December 2020

Engineered proteins as tools to understand ubiquitin signaling

Lin Hui Chang

Engineered Proteins As Tools To Understand Ubiquitin Signaling

A Dissertation Presented

by

LIN HUI CHANG

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2020

Chemistry

Engineered Proteins As Tools To Understand Ubiquitin Signaling

A Dissertation Presented

by

LIN HUI CHANG

Approved as to style and content by:

_____

Eric R. Strieter, Chair

_____

Lynmarie K. Thompson, Member

_____

Min Chen, Member

_____

Scott C. Garman, Outside Member

_____

Ricardo B. Metz, Department Head

Chemistry Department

ACKNOWLEDGEMENT

I would like to thank my advisor, Eric Strieter, for his guidance and encouragement. I still remember vividly that after my transfer to University of Massachusetts-Amherst, that was a time when I was struggling with my research; he has asked me to focus on yeast display and has accompanied me to seek advice from other faculty member in the field. Eric is very passion in science and will always come into the lab with different ideas, sometimes they could be weird, but most of the times, his ideas are thought-provoking. His attitude in science will continue motivate me to strive harder in my future career.

I would also like to thank my committee members, Prof. Lynmarie Thompson, Prof. Min Chen and Prof. Scott Garman for their advices in my research. I really appreciate working with the core facilities manager, especially for Dr. Amy Burnside, Dr. Lizz Bartlett, Dr. Ravi Ranjan, Dr. James Chambers, and Dr. Steve Eyles for being always available every time I have questions with the instruments or experiments. I would like to especially thank Prof. Sarah Moore from Smith College and her student, Allison Sirois for helping me with yeast display.

I am fortunate to work with my group members. They have taught me how to perform experiments and helped me with my writing and presentation for my exam. Robert, Ambar, Sean, Kirandeep and Jiale, whom I know from Madison. Because of them, I have the courage to move to Amherst to challenge the new environment. I would like to thank the group members in Amherst, Sandor,

Yanfeng, Heather, Fangying, Rehab and Rakesh. Thank you all for the numerous supports, and sharing the ideas and materials with me.

Also, I would like to thank the Chemistry Department. DV, jms, and Chandra who help me with the administrative documents. Vicki and Kris for the helps with the issues regarding purchasing. Marv, Bob, and Ryan, for fixing our instruments and help us every time something wrong in the lab. John, Gary, and Nikki who come into rescue when my computer decided to “have an attitude”. Dennis who is so friendly, and a trip to the stockroom is always a good short break from the experiments.

Last but not least, I am very grateful for families and friends. Dad and mum, who always encourage me every time I am doubtful of myself and my sisters, Lin Zhi, Lin June and Lin-li who support me to pursuit my graduate study. My relatives in New York, who is like a second family to me and make sure I adapt to the life in USA. Also, I would like to thank to my friends, especially Ying Ji, Kong, and 7th floor members for all the morale support throughout my graduate career.

ABSTRACT

ENGINEERED PROTEINS AS TOOLS TO UNDERSTAND

UBIQUITIN SIGNALING

SEPTEMBER 2020

LIN HUI CHANG, B. S., UNIVERSITY OF TOKYO

M.S., UNIVERSITY OF WISCONSIN-MADISON

PH. D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by Professor Eric R. Strieter

Ubiquitin is a 76 amino acids protein that is evolutionary conserved in eukaryotes. It is an important signaling molecule in a plethora biological events, such as protein degradation, DNA damage response, and transcription. This thesis aims to develop engineered protein as a tool to study ubiquitin signaling. Through targeted mutagenesis and directed evolution, a deubiquitinase is reprogrammed into a transamidase, which lead to the generation of ubiquitin-protein conjugates with discrete ubiquitin linkages through auto-ubiquitination. These ubiquitin-protein conjugates could be used as a model substrate to profile their interaction of different ubiquitin interacting proteins. In addition, using directed evolution and deep sequencing, a nanobody-based binder targeting a deubiquitinase is engineered. This nanobody could be utilized as an intracellular probe to study the regulation of the deubiquitinase with great spatial and temporal control.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Ubiquitination machinery

In the late 1970s, a heat-stable protein, named as ATP-dependent proteolysis factor 1 (APF-1) was elucidated to play a crucial role in the ATP-dependent proteolysis; it is further demonstrated that single or multiple APF-1s are conjugated to the substrate to activate the proteolysis of the substrate.[1,2,3] APF-1 is later shown to be ubiquitin and the discovery of the ubiquitin dependent proteolysis has been awarded with 2004 Nobel Prize in Chemistry.[4]

Ubiquitin is a 76 amino acids protein and is evolutionally conserved in eukaryotes. Since the discovery of the role of ubiquitin as the signaling molecule in proteolysis, ubiquitin has been shown to be involve in a plethora of biological events. This process, known as ubiquitination, is achieved through the formation of isopeptide or peptide bond between the C terminal of ubiquitin with the α-amino group of the N terminus or ε-amino group of the lysine of the substrate. This conjugation is an ATP dependent enzymatic cascade involving ubiquitin activating enzyme (E1), ubiquitin conjugating enzyme (E2) and ubiquitin ligase (E3).[5] These enzymes are referred as the ubiquitin writers. Their actions are reversible, and the removal of ubiquitin, or the ubiquitin editing are performed by deubiquitinase (**Figure 1-1**).[6]

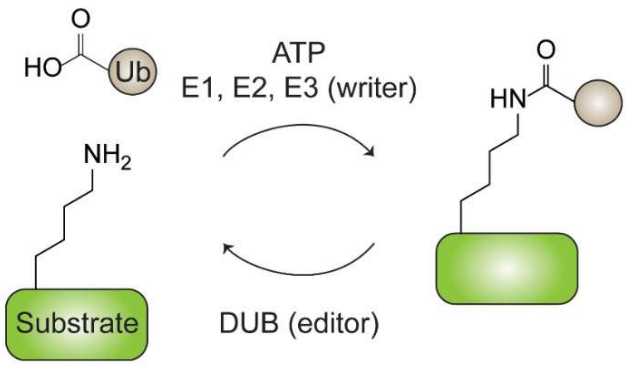

Figure 1-1: Ubiquitination machinery

In addition, ubiquitin contains eight residues with primary amines; other than the α-amino group on the Met1, ubiquitin has seven lysine which are Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 and Lys63 (**Figure 1-2A**). Hence, a single ubiquitin can be further connected to another ubiquitin at these residues to form different polyubiquitin. Here, the lysine donating moiety is referred as the proximal ubiquitin and the lysine accepting molecule is the distal ubiquitin. Here, homotypic chain describes the ubiquitin chain which is formed from the conjugation at the same lysine, while the heterotypic chain refers to the ubiquitin chain containing different linkages. Especially for heterotypic chain, it could be further categorized into mixed and branched polyubiquitin (**Figure 1-2B**). It has been reported that all kind of ubiquitin linkages present *in cellulo*, although 60% of the total cellular ubiquitin is estimated to be in the form of monoubiquitination compare to 11% for ubiquitin chain.[7,8,9,10]

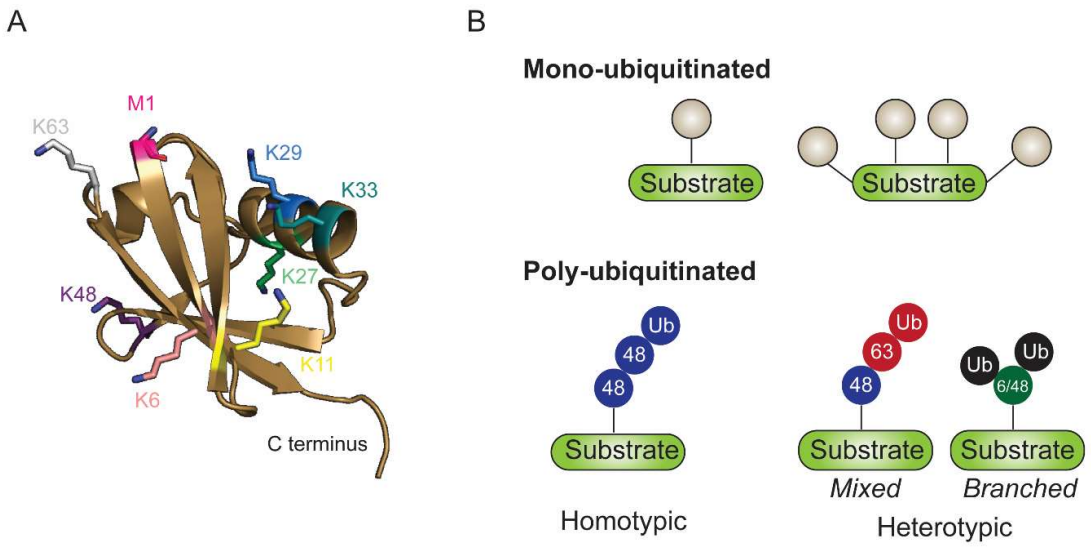

Figure 1-2: Ubiquitin and its conjugates

(**A**) Crystal structure of ubiquitin (PDB ID: 1UBQ). The Met1 and its seven lysine residues are highlighted in color. (**B**) Examples of ubiquitin conjugates.

How are different ubiquitin conjugates recognized by their interacting proteins?

Different ubiquitin conjugates have been reported to elicit distinct cellular responses. For example, mono-ubiquitinated histone has been associated with transcriptional regulation, Lys48 linked ubiquitinated substrates are subjected to proteasomal degradation and Lys63 linked ubiquitin conjugates are reported to participate in DNA damage response and NF-κB signaling pathway.[11,12,13,14] This prompts a perception that the different ubiquitin conjugates adopt different topologies and hence, are recognized by certain ubiquitin interacting proteins, such as the deubiquitinases or ubiquitin binding domains (also referred as the readers).

Extensive structural and biochemical studies have demonstrated that ubiquitin conjugates use distinct surface regions to interact with their interacting

proteins. The main interaction surfaces of ubiquitin are at the hydrophobic surfaces center at Phe4, Ile36 and Ile44 of the ubiquitin.[15,16] In addition, the C terminal of ubiquitin which is involved in conjugation is another important interacting surfaces. For example, the zinc finger ubiquitin binding domain (ZnF UBP) of USP5 develops a deep binding cleft to recognize the C terminal GlyGly motif of ubiquitin.[17] The conformation of ubiquitin chain with different linkages are determined by these surfaces too. Majority of the Lys48 linked ubiquitin dimer adopts a 'closed' conformation and interacts through the Ile44 patches; the UBA2 domain of the hHR23A, which is a proteasomal shuttling factor is shown to sandwiched between the Lys48 linked ubiquitin dimer and interact with both ubiquitin moieties at the same hydrophobic surface.[18]

Amounting evidence has suggested that the activities of deubiquitinases are influenced by the substrate which is ubiquitinated too. While the catalytic domains of these deubiquitinases are responsible to recognize ubiquitin, the specificity of the deubiquitinases are benefited from the interaction between the external domains with the substrate. USP7 is reported to recruit substrates by its N terminal TRAF domain and preferentially remove ubiquitin chains en bloc from its targets, such as p53 and PCNA, but not breaking down the ubiquitin chains.[19,20] These studies have underscored that other than the ubiquitin linkages, it is crucial to investigate the substrate in order to understand the interactions of ubiquitin conjugates with their editors or readers. *This is non-trivial, as it remains a great challenge to produce defined ubiquitin protein*

conjugates. Numerous efforts have been devoted to innovating the generation of ubiquitin protein conjugates.

How the specificities of ubiquitin interacting protein are regulated?

Ubiquitin conjugates involve in various signaling pathways inside the cell, and aberrant interactions of the ubiquitin conjugates with their interacting proteins, such as the deubiquitinases and ubiquitin binding domains, are detrimental. In fact, anomalous activities of these ubiquitin interacting proteins had been elucidated to lead to numerous diseases such as cancer and neurological diseases. The overexpression of RPN13, which is a proteasomal ubiquitin receptor has a poor prognosis in ovarian cancer;[21] and a E7A mutation of UCHL1, which is a deubiquitinase, has reported to have reduced activity, and the loss of function leads to an early-onset progressive neurodegenerative syndrome.[22] Hence, understanding how the ubiquitin interacting proteins are regulated could give insight into the development of novel therapeutic intervention. Generally, these deubiquitinases and ubiquitin binding domains are regulated at their expression or abundance level, post translational modification, and recruitment to macromolecular complexes.

Regulation of the proteasome associated ubiquitin receptors and deubiquitinases

Here, we would use 26S proteasome as an example to explain the regulation of ubiquitin editor and readers that are involved in the ubiquitin proteasome system. 26S Proteasome is a 2.6 MDa complex consists of around

30 subunits. These proteasome subunits are categorized into 20S CP (core particle), which consists of all the subunits in the proteolytic chamber and 19S RP (regulatory particle, also known as PA700), containing the ATPases and non-ATPase regulatory units, including the proteasomal ubiquitin receptors, RPN10 and RPN13 and proteasomal deubiquitinases, RPN11, UCH37 and USP14 (**Figure 1-3A**). [23] As the main protein degradation machinery in eukaryotes, proteasome degrades the protein through the ubiquitin proteasome system (UPS). In this system, the proteasome is recruited to the ubiquitinated substrate through its ubiquitin receptors. The proteasome associated deubiquitinases will break down the ubiquitin chain or remove the ubiquitin from the substrate. This is followed by the unfolding of substrate by the AAA+ ATPases. In addition to the unfolding, the ATPases also trigger the opening of a gated access channel and facilitate the translocation of the substrate into the peptidase chamber for proteolysis (**Figure 1-3B**).[24] Proteasome has been a promising therapeutic target in oncology therapeutic. Bortezomib, under the brand name Velcade®, which inhibit the peptidases, had been approved FDA to treat patients with multiple myeloma with patients achieve complete remission.[25] Currently, proteasome associated ubiquitin interacting proteins also gain great interest as a pharmacological target due to their well-developed binding pockets. Numerous studies are now underway to study their role in cellular physiology.

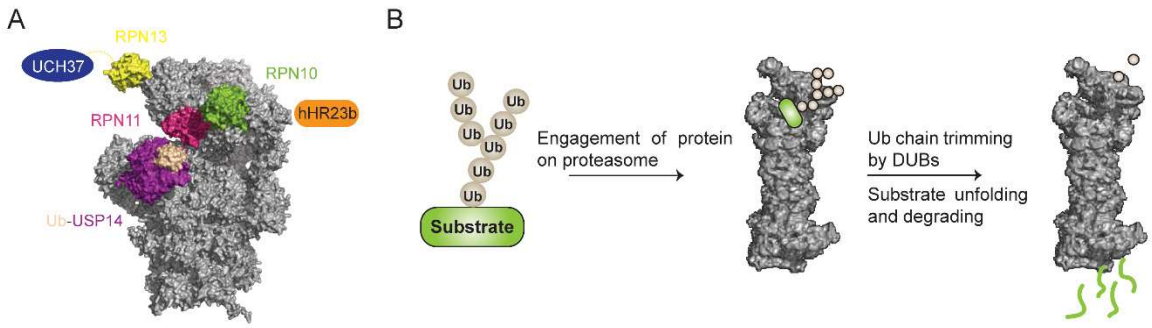

Figure 1-3: Proteasome

(**A**) Representative cryo-EM structure of proteasome (PDB ID: 5A5B) with proteasome associated deubiquitinases and ubiquitin readers being highlighted in color. (**B**) Ubiquitin-proteasome system.

Expression level

Expression of the proteasome subunits are subjected to dynamic regulation to maintain protein homeostasis.[26,27] In steady state, the basal expression proteasome subunits are controlled by the transcription factor NF-Y through the consensus sequence at the promoter regions.[28] Nonetheless, it has demonstrated that certain proteasome subunits are overexpressed in solid tumors; they are RPN1 and RPN13, which are the proteasomal subunits recognizing linear and branched Lys48 linked ubiquitin chains.[29,30] On the other hand, it has been elucidated that while the challenge with proteasome inhibitors, such as bortezomib or carfilzomib, would increase the transcription level of most proteasome subunits by the transcription factor Nrf1, the reduction of 19S RP subunits, including RPN11 and RPN13 would convey resistance to those proteasome inhibitors across a broad spectrum of cancer cell lines.[31,32,33] These findings highlight the importance to investigate the differential expression levels

of the proteasome associated ubiquitin receptors and deubiquitinases in the development of effective therapeutics targeting the ubiquitin proteasome system.

Post translational modification

In addition, the proteasome associated ubiquitin readers and deubiquitinases could be regulated through post translational modifications, such as phosphorylation and ubiquitination. These modifications on these subunits influence the activity of proteasome either (i) by modulating their catalytic activities or (ii) by interacting differently with the macromolecule complex or with ubiquitin conjugates. For example, the catalytic activity of USP14 changes upon post translational modification. USP14 is shown to be phosphorylated at Ser432 by Akt and it is speculated that the negatively charged phosphate group induces a rearrangement of the autoinhibitory BL2 loop of USP14, hence, activates its deubiquitinating activity *in vitro* and *in vivo*. It was further demonstrated that USP14 KO cells with exogenous USP14 containing phosphomimetic mutations has decreased global proteasomal degradation.[34]

Post translational modification could positively and negatively regulates the interactions of the proteasomal subunits with proteasome. hHR23B, which is a proteasomal shuttling factor that transports ubiquitinated substrate to proteasome, could be phosphorylated at its UBL domain, which is the domain that bind to UBL receptors of proteasome. While this phosphorylation weakens the interaction between hHR23B and proteasome in *Saccharomyces cerevisiae*, the role of the phosphorylation of hHR23B in mammalian cells is yet to be

studied.[35] There are examples where phosphorylation events would improve the binding with the proteasome. The phosphorylated RPN2, which is a proteasome scaffolding protein at Tyr950 has a higher binding affinity towards RPN13 *in vitro*.[36]

On the other hand, ubiquitination of the proteasome associated subunits has been reported to modify their interactions with the ubiquitin conjugates. In yeast, RPN10 is shown to be monoubiquitinated. It was demonstrated that RPN10 mutant which could not be ubiquitinated could not complement the *rad23* and *rpn10* null yeast cell, suggesting a possible role of the monoubiquitination species in important for normal cellular physiology. Nonetheless, the monoubiquitinated RPN10 shows lower binding affinity towards ubiquitin chains, possibly via the inhibition at the ubiquitin interacting motif (UIM) of RPN10; this has resulted in the lower rescue efficiency in the yeast with defective proteasome lacking RPN10 and RAD23.[37] In addition, RPN13 is significantly autoubiquitinated upon proteotoxic insult. *In vitro*, ubiquitination of RPN13 inhibits its binding affinity of proteasome towards ubiquitinated substrate, and hence decrease the capacity of proteasome to degrade the substrate. Although the functional consequences of the autoubiquitinated RPN13 in a cellular context remains to be elusive, it is suggested this modification could involve in a negative feedback loop during proteotoxic stress.[38]

Recruitment to proteasome

Proteasome itself is a dynamic macromolecule complex; in general, the association between 19S RP and 20S CP to assembly the 26S proteasome is labile and sensitive to the environment. Under oxidative stress, the 19S RP would dissociate to facilitate the degradation of the oxidized protein by 20S CP in an ATP and ubiquitin independent pathway.[39,40] In an inflammatory response, the cytokines such as TNF-α and IFN-γ trigger the swapping of the β1, β2 and β5 subunits of 20S CP to β1i, β2i and β5i subunits. The 26S proteasome would also disassemble with the 20S CP assemblies with other activators such as PA28αβ to generate immunoproteasome.[26]

Interestingly, little is known about the assembly and the dynamics of the 19S RP, although studies have elucidated that the several components of 19S RP are dynamic interactors of the 19S RP. A quantitative cross-linking mass spectrometry has demonstrated for the purified proteasome, the RPN13 and UCH37, which is recruited to proteasome through RPN13, are at sub-stoichiometry level in various mammalian cell lines. The same observation is for USP14.[41,42] It was further shown that these interactions are dynamic, and these proteasome associated subunits are readily swapped between the proteasome bound and free state.[43] These observations are interesting as it has been suggested that these deubiquitinases may have difference ubiquitin editing preferences: UCH37 preferentially hydrolyzes Lys48 linked branched ubiquitin chain, while USP14 removes ubiquitin chain *en bloc*.[44,45] This is in contrast to the RPN11, which is the constitutive proteasome associated deubiquitinase that

remove the ubiquitin chain *en bloc* by hydrolyzing the isopeptide linkage between the proximal ubiquitin and substrate upon the translocation of substrate into the 20S CP.[46] Hence, it leads to the hypothesis that proteasome with different composition of deubiquitinases may have different deubiquitination pattern.

In addition, both USP14 and UCH37 are suggested to regulate proteasome activity towards distinct substrates.[47,48] Through *in vitro* complementation test, proteasome with WT USP14 shows decreased degradation compared to proteasome with catalytically inactive USP14; while proteasome depleted of both USP14 and UCH37 also has lower degradation efficiency compare to proteasome lack of USP14 only.[49,44] This lead to the speculation that the dynamic recruitment of USP14 and UCH37 to the proteasome might generate proteasome with different degradation capacity in a cellular context. The question is, what is the trigger of the (dis)incorporation of these deubiquitinases? What are the functional consequences of a proteasome lack of USP14 or/and UCH37? *Advanced intracellular probes for these ubiquitin readers or editors should be developed to monitor their interactions with proteasome spatially and temporally.*

In short, findings suggest that these proteasome-associated ubiquitin readers and editors play a crucial role in determining the ubiquitin dependent proteasomal degradation. Hence, they are under dynamic regulations at multi-level, possibly to fine tune the proteasome degradation capacity under different cellular conditions to promote cellular survival.

Protein engineering in ubiquitin system

In the previous section, we highlighted two emerging questions in the study of ubiquitin signaling: (i) the generation of defined ubiquitin protein conjugates to understand their interactions with the ubiquitin editors or readers and (ii) the generation of intracellular probe for ubiquitin interacting proteins. In this thesis, we would like to tackle the questions through protein engineering. Protein engineering describes the process to develop proteins or enzymes with desirable functions and has been widely applied in both academic and industrial, and the methodologies and techniques developed in protein engineering are awarded with 2018 Nobel Prize in Chemistry. The research in the ubiquitin field also benefited from protein engineering, and numerous next generation probes targeting the ubiquitin system has been engineered.[50]

Rational design and directed evolution

The two mains methodologies in protein engineering are rational design and directed evolution. In rational design, prior knowledge, such as the structure, sequence alignment, and function or mechanism of a protein or its homologs, is critical to guide the engineer of a protein. The precise amino acids of the engineered protein are mutated based on site-directed mutagenesis, and the mutants are characterized to determine whether they carry the desired properties. Rational design has deepened our understanding about the role of amino acids in determining the binding and catalysis of a protein while laying out the foundation for the computational design of *de novo* protein. Nonetheless,

rational design requires a tedious characterization of protein mutants and limit the number of mutants to be tested. Also, it could be impractical if nothing is known about the protein to be designed.[51]

On the other hand, directed evolution could circumvent the lack of knowledge of the structure and function of the protein. The booming of the directed evolution is benefited from the advances in gene diversification, high throughput screening methods through various display technologies, and deep sequencing. In principle, the protein of interest is subjected to mutagenesis, either through targeted or random mutagenesis, to generate an ensemble of mutants. This ensemble, also referred as library, is challenged with an artificial selection and screening to enrich for the mutants with improved function, while the mutants with poor outcome are discarded (**Figure 1-4**).[52] The power of directed evolution is that a library with large amount of mutants, with 10^6 – 10^{13} could be generated in the laboratory which would be selected or screened in a high throughput fashion, and the enriched mutants could be further identified through next generation sequencing to give insight into the beneficial mutations at amino acid level.

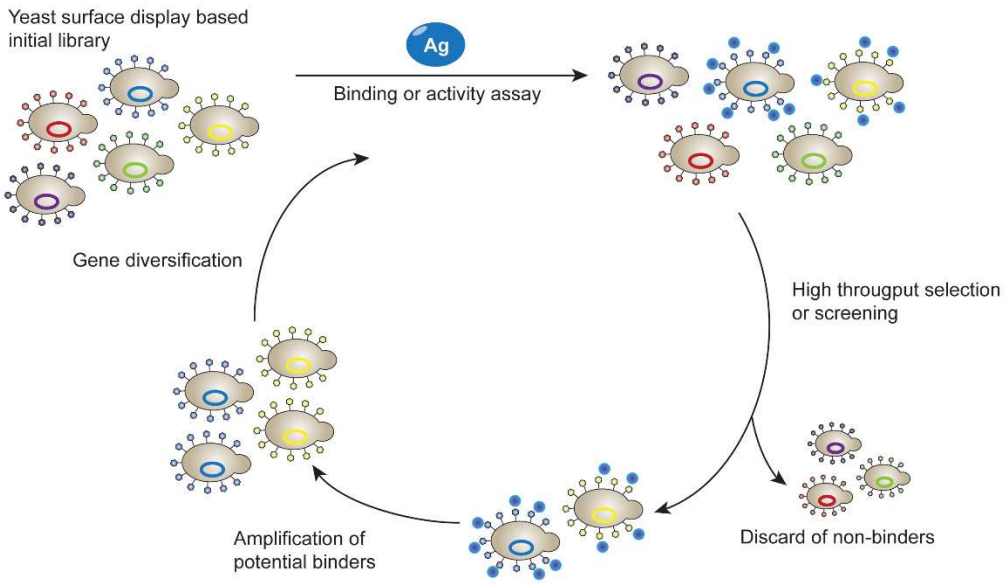

Figure 1-4: Directed evolution

Scheme of directed evolution exemplified by yeast surface display.

While the traditional way to generate a library is through error prone PCR reaction to introduce random mutation on the gene of interest, the improvement in synthesis of DNA oligos has further expanded the protein sequence space. For example, DNA oligos with degenerate codons allow saturation mutagenesis at certain regions of the protein. In addition, homologous recombination, such as DNA shuffling, permits the combination of enriched gene which may yield a synergistic effect.[53]

Various display methods have been developed to facilitate the directed evolution of a protein. These display methods linked the genotype of a to-be-designed protein with its phenotype to enable the identification of the isolated clones through the high throughput selection or screening strategies. The proteins could be (i) displayed on the surface of phage, bacterial, yeast and

mammalian cells, (ii) fused to molecules such as ribosome or mRNA.[54,55,56,57,58,59] Recently, microfluidics technology has been applied in the *in vitro* compartmentalization of gene of interest together with a transcription and translational machinery, providing an alternative to establish the phenotype-genotype linkage necessary for directed evolution.[60] These different methods, which determine the size of the library and the suitable selection and screening strategies have their pros and cons. For example, the transformation efficiency of yeast surface display has limited the size of library to be $10^7 – 10^9$, which are smaller compare to size of library generated by phage display at the order of 10^{10}. Nonetheless, yeast surface display could be coupled with fluorescent activated cell sorting (FACS) with outstanding advantages: each clone could be scrutinized in FACS, hence it is efficient for positive and negative screenings, and the multicolor capabilities of FACS enable the quantitative control through normalization of display level.[61]

In addition, the application of next generation sequencing has revolutionized the directed evolution of protein, especially in the engineering of a protein binder. The increased sampling depth, which is in the order of 10^6 and the quantitative assessment of the isolated clones offered by the high throughput sequencing platform have led to the discovery of binders with improved binding kinetics.[62] In addition, next generation sequencing is capable to detect clones with marginal improvement in binding; these small beneficial mutations could be synergistic and essential to overcome the “fitness valley” in the protein engineering landscape.[63]

Engineered protein as tools to study ubiquitin system

Despite their differences, rational design and directed evolution are complementary. Many research labs, including in the ubiquitin field have combined both techniques to generate engineered protein as tools to study the ubiquitin system. For example, E3 ligases have been engineered to catalyze the generation of certain conjugate. Here, a modular human E3 ligase, CHIP (carboxyl terminus of Hsc70-interacting protein) is reprogrammed to recruit specific substrate to the E3 ligase to induce the ubiquitination of the substrates. The recruitment is achieved by fusing the CHIP to a substrate specific single-chain variable fragment (scFv) or a fibronectin type III domain monobody. It was further demonstrated that the engineered CHIP could be expressed intracellularly to ubiquitinate and to further degrade the targeted substrate.[64] This study revealed that the induced ubiquitination through proximity effect by a modular ubiquitin ligase is possible. Nonetheless, study has shown this strategy yield a mixture of ubiquitin protein conjugate with ubiquitin chain of different length and linkages. It remains a great challenge to prepare ubiquitin protein conjugate with distinct ubiquitin chain architecture.

Different engineered proteins have been developed to facilitate the study of ubiquitin chains. The K11/K48-bispecific antibody targeting heterotypic Lys11- and Lys48 linked ubiquitin chain is designed by the knobs-into-holes technology to heterodimerize the monospecific Lys11 or Lys48 antibodies. The antibody, which displays higher binding affinity towards heterotypic chain than the homotypic K11- or K48 linked polyubiquitin, had been applied successfully in

Western blot, immunoprecipitation and immunofluorescence.[65] In addition, affimers targeting Lys6 linked and Lys33 linked ubiquitin chain has been engineered. Affimer is a synthetic binding protein and the linkage specific affimers have given insight into the distinct conformation of the ubiquitin linkages.[66]

Other than being the tools to recognize ubiquitin conjugates, engineered proteins targeting the deubiquitinases and ubiquitin readers had been designed. By using ubiquitin as scaffold and phage display, the Sidhu's lab generated different ubiquitin variants (UbV) targeting proteins in the ubiquitin system, for example E2 and E3 ubiquitinating enzyme, deubiquitinases from the USP families, and ubiquitin binding domain like ubiquitin interacting motif (UIM). Structural studies elucidated that all the ubiquitin variants target the same active site or binding site as the wildtype ubiquitin. Hence, these ubiquitin variants could be used as the intracellular probes which modulate the activities, either inhibition or activation, of their targets.[67,68] These findings also indicated that due to the structurally conserved ubiquitin binding site shared by the ubiquitin writers, editors, and readers, it is inevitable that the ubiquitin derived protein would occupy the same site. Nonetheless, since different ubiquitin binding sites have low sequence consensus, it is possible to fine tune the specificity of the ubiquitin variants by exploiting the interactions of ubiquitin variants with the non-conserved sequences.

Ubiquitin variants which bind the ubiquitin binding sites are ideal to be developed as the probes which modulate the activities of their target.

Nonetheless, applications such as intracellular imaging require the probe to have minimal perturbation on the target, which is inapplicable for ubiquitin variants. On the other hand, there are examples of protein binders with non-ubiquitin scaffold targeting the ubiquitin readers. DARPin (designed ankyrin-repeat proteins) targeting the CC2-LZ domain of NEMO has been designed through ribosome display. NEMO is a regulatory component in the NF-κB signaling. It dimerizes to bind to linear ubiquitin chain and both dimerization and ubiquitin interactions involve the CC2-LZ domain. Here, this DARPin was utilized as a chaperon to assist the crystallization of the dimeric CC2-LZ domain and as an intracellular inhibitor of the NF-κB signaling.[69,70] Crystal structure suggested the DARPin also interacts with the NEMO at the same interaction surfaces as ubiquitin as well. This is not surprising, as the binding site for the substrate, that is the ubiquitin, is a well evolved binding cleft.

In short, engineered proteins have play an active role in helping the scientific community in understanding the ubiquitin system. The well-developed methodologies and techniques in protein engineering have facilitated the generation of tools to study the ubiquitin signaling.

Forward to thesis

Studies have shown that substrates with different ubiquitin topologies and chain lengths would have different biological outcomes. Nonetheless, the major challenge in understanding the interactions of the ubiquitin conjugates and their interacting proteins is to generate ubiquitin protein conjugates with defined

ubiquitin linkages. Chapter 2 describes a facile method to generate a ubiquitinated protein which is ubiquitinated site specifically with distinct ubiquitin chain. This is achieved through the reprogramming of a deubiquitinase, Yuh1, into an auto-ubiquitinating ligase through a combination of targeted mutagenesis and yeast surface display for directed evolution. We further demonstrated that this ubiquitin protein conjugate could be a model substrate to profile DUBs activities.

On the other hand, understanding how the ubiquitin editors or readers are regulated are crucial to obtain a complete picture of the ubiquitin system. Here, small protein binders are a powerful tool to study target protein *in vitro*, *in cellulo*, and *in vivo*. In chapter 3, we describe the generation of a nanobody with high binding affinity towards a proteasome-associated deubiquitinase, UCH37, through directed evolution using yeast surface display. We further confirmed the interaction of the nanobody with proteasome-associated UCH37 in a cellular context. This suggests the nanobody could be developed as the intracellular probe for proteasomal UCH37.

As one of the proteasome-associated deubiquitinase, UCH37 is responsible to specifically hydrolyze the Lys48 linked branched ubiquitin chains. It was shown that UCH37 interacts with proteasome in a dynamic fashion. What is the physiological relevance of the transient interaction of the UCH37 with the proteasome? We proposed to use the nanobody specific towards proteasome associated UCH37 as an intracellular imaging probe to investigate the recruitment of UCH37 to proteasome. In Chapter 4, we first established that the

discovered nanobody does not interfere with the catalytic activities of the proteasome-associated UCH37 and could be developed as a probe for live-cell imaging.

References

1. Ciehanover, A., Hod, Y. & Hershko, A. A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem. Biophys. Res. Commun.* **81**, 1100–1105 (1978).
2. Hershko, A., Ciechanover, A. & Rose, I. A. Resolution of the ATP dependent proteolytic system from reticulocytes: A component that interacts with ATP. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 3107–3110 (1979).
3. Hershko, A., Ciechanover, A., Heller, H., Haas, A. L. & Rose, I. A. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 1783–1786 (1980).
4. Wilkinson, K. D., Urban, M. K. & Haas, A. L. Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *J. Biol. Chem.* **255**, 7529–7532 (1980).
5. Pickart, C. M. & Eddins, M. J. Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta - Mol. Cell Res.* **1695**, 55–72 (2004).
6. Reyes-Turcu, F. E., Ventii, K. H. & Wilkinson, K. D. Regulation and Cellular Roles of Ubiquitin-Specific Deubiquitinating Enzymes. *Annu. Rev. Biochem.* **78**, 363–397 (2009).
7. Xu, P. *et al.* Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. *Cell* **137**, 133–145 (2009).
8. Crowe, S. O., Rana, A. S. J. B., Deol, K. K., Ge, Y. & Strieter, E. R. Ubiquitin Chain Enrichment Middle-Down Mass Spectrometry Enables Characterization of Branched Ubiquitin Chains in Cellulo. *Anal. Chem.* **89**, 4428–4434 (2017).
9. Clague, M. J., Heride, C. & Urbé, S. The demographics of the ubiquitin system. *Trends Cell Biol.* **25**, 417–426 (2015).
10. Kaiser, S. E. *et al.* Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nat. Methods* **8**, 691–696 (2011).

11. Cao, J. & Yan, Q. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Front. Oncol.* **2**, 1–9 (2012).

12. Pickart, C. M. Targeting of substrates to the 26S proteasome. *FASEB J.* **11**, 1055–1066 (1997).

13. Uckelmann, M. & Sixma, T. K. Histone ubiquitination in the DNA damage response. *DNA Repair (Amst).* **56**, 92–101 (2017).

14. Xia, Z. P. *et al.* Direct activation of protein kinases by unanchored polyubiquitin chains. *Nature* **461**, 114–119 (2009).

15. Komander, D. & Rape, M. The Ubiquitin Code. *Annu. Rev. Biochem.* **81**, 203–229 (2012).

16. Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M. & Hicke, L. Distinct Functional Surface Regions on Ubiquitin. *J. Biol. Chem.* **276**, 30483–30489 (2001).

17. Reyes-Turcu, F. E. *et al.* The Ubiquitin Binding Domain ZnF UBP Recognizes the C-Terminal Diglycine Motif of Unanchored Ubiquitin. *Cell* **124**, 1197–1208 (2006).

18. Varadan, R., Assfalg, M., Raasi, S., Pickart, C. & Fushman, D. Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain. *Mol. Cell* **18**, 687–698 (2005).

19. Sahtoe, D. D. & Sixma, T. K. Layers of DUB regulation. *Trends Biochem. Sci.* **40**, 456–467 (2015).

20. Masuda, Y., Kanao, R., Kawai, H., Kukimoto, I. & Masutani, C. Preferential digestion of PCNA-ubiquitin and p53-ubiquitin linkages by USP7 to remove polyubiquitin chains from substrates. *J. Biol. Chem.* **294**, 4177–4187 (2019).

21. Fejzo, M. S. *et al.* Comprehensive Analysis of 20q13 Genes in Ovarian Cancer Identifies ADRM1 as Amplification Target. *Genes. Chromosomes Cancer* **47**, 873–883 (2008).

22. Bilguvar, K. *et al.* Recessive loss of function of the neuronal ubiquitin hydrolase UCHL1 leads to early-onset progressive neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 3489–3494 (2013).

23. Lander, G. C. *et al.* Complete subunit architecture of the proteasome regulatory particle. *Nature* **482**, 186–191 (2012).

24. Tanaka, K. The proteasome: Overview of structure and functions. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* **85**, 12–36 (2009).

25. Paubelle, E. *et al.* Complete remission with bortezomib on plasmocytomas in an end-stage patient with refractory multiple myeloma who failed all other therapies including hematopoietic stem cell transplantation: Possible enhancement of graft-vs-tumor effect [15]. *Leukemia* **19**, 1702–1704 (2005).

26. Motosugi, R. & Murata, S. Dynamic regulation of proteasome expression. *Front. Mol. Biosci.* **6**, 4–11 (2019).

27. Voutsadakis, I. A. Proteasome expression and activity in cancer and cancer stem cells. *Tumor Biol.* **39**, 1–17 (2017).

28. Xu, H. *et al.* The CCAAT box-binding transcription factor NF-Y regulates basal expression of human proteasome genes. *Biochim. Biophys. Acta - Mol. Cell Res.* **1823**, 818–825 (2012).

29. Pilarsky, C., Wenzig, M., Specht, T., Saeger, H. D. & Grützmann, R. Identification and validation of commonly overexpressed genes in solid tumors by comparison of microarray data. *Neoplasia* **6**, 744–750 (2004).

30. Boughton, A. J., Krueger, S. & Fushman, D. Branching via K11 and K48 Bestows Ubiquitin Chains with a Unique Interdomain Interface and Enhanced Affinity for Proteasomal Subunit Rpn1. *Structure* **28**, 29-43.e6 (2020).

31. Radhakrishnan, S. K. *et al.* Transcription Factor Nrf1 Mediates the Proteasome Recovery Pathway after Proteasome Inhibition in Mammalian Cells. *Mol. Cell* **38**, 17–28 (2010).

32. Tsvetkov, P. *et al.* Compromising the 19S proteasome complex protects cells from reduced flux through the proteasome. *Elife* **4**, 1–22 (2015).

33. Acosta-Alvear, D. *et al.* Paradoxical resistance of multiple myeloma to proteasome inhibitors by decreased levels of 19S proteasomal subunits. *Elife* **4**, 1–19 (2015).

34. Xu, D. *et al.* Phosphorylation and activation of ubiquitin-specific protease-14 by Akt regulates the ubiquitin-proteasome system. *Elife* **4**, 1–16 (2015).

35. Liang, R. Y. *et al.* Rad23 interaction with the proteasome is regulated by phosphorylation of its ubiquitin-like (UbL) domain. *J. Mol. Biol.* **426**, 4049–4060 (2014).

36. Hemmis, C. W., Heard, S. C. & Hill, C. P. Phosphorylation of Tyr-950 in the proteasome scaffolding protein RPN2 modulates its interaction with the ubiquitin receptor RPN13. *J. Biol. Chem.* **294**, 9659–9665 (2019).

37. Isasa, M. *et al.* Monoubiquitination of RPN10 Regulates Substrate Recruitment to the Proteasome. *Mol. Cell* **38**, 733–745 (2010).

38. Besche, H. C. *et al.* Autoubiquitination of the 26S Proteasome on Rpn13 Regulates Breakdown of Ubiquitin Conjugates. *EMBO J.* **33**, 1159–1176 (2014).

39. Wang, X., Yen, J., Kaiser, P. & Huang, L. Regulation of the 26S proteasome complex during oxidative stress. *Sci. Signal.* **3**, ra88–ra88 (2010).

40. Aiken, C. T., Kaake, R. M., Wang, X. & Huang, L. Oxidative Stress-Mediated Regulation of Proteasome Complexes. *Mol. Cell. Proteomics* **10**, R110.006924 (2011).

41. Fabre, B. *et al.* Label-free quantitative proteomics reveals the dynamics of proteasome complexes composition and stoichiometry in a wide range of human cell lines. *J. Proteome Res.* **13**, 3027–3037 (2014).

42. Fabre, B. *et al.* Subcellular distribution and dynamics of active proteasome complexes unraveled by a workflow combining in vivo complex cross-linking and quantitative proteomics. *Mol. Cell. Proteomics* **12**, 687–699 (2013).

43. Wang, X. & Huang, L. Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. *Mol. Cell. Proteomics* **7**, 46–57 (2008).

44. Deol, K. K., Crowe, S. O., Du, J., Bisbee, H. & Strieter, E. R. Proteasome-Bound UCH37 Debranches Ubiquitin Chains to Promote Degradation 1 2. doi:10.1101/2020.02.21.960088.

45. Lee, B.-H. *et al.* USP14 deubiquitinates proteasome-bound substrates that are ubiquitinated at multiple sites. *Nature* **532**, 398–401 (2016).

46. de Poot, S. A. H., Tian, G. & Finley, D. Meddling with Fate: The Proteasomal Deubiquitinating Enzymes. *J. Mol. Biol.* **429**, 3525–3545 (2017).

47. Chadchankar, J. *et al.* Inactive USP14 and inactive UCHL5 cause accumulation of distinct ubiquitinated proteins in mammalian cells. *PLoS One* **14**, 1–23 (2019).

48. Liu, B. *et al.* Proteome-wide analysis of USP14 substrates revealed its role in hepatosteatosis via stabilization of FASN. *Nat. Commun.* **9**, 1–12 (2018).

49. Lee, B. H. *et al.* Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* **467**, 179–184 (2010).

50. Zhao, B. *et al.* Protein engineering in the ubiquitin system: Tools for discovery and beyond. *Pharmacol. Rev.* **72**, 380–413 (2020).

51. Chen, R. Enzyme engineering: Rational redesign versus directed evolution. *Trends Biotechnol.* **19**, 13–14 (2001).

52. Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **10**, 866–876 (2009).

53. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nat. Publ. Gr.* **16**, 379–394 (2015).

54. Smith, G. P. & Petrenko, V. A. Phage display. *Chem. Rev.* **97**, 391–410 (1997).

55. Daugherty, P. S. Protein engineering with bacterial display. *Curr. Opin. Struct. Biol.* **17**, 474–480 (2007).

56. Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **15**, 553–557 (1997).

57. Akamatsu, Y., Pakabunto, K., Xu, Z., Zhang, Y. & Tsurushita, N. Whole IgG surface display on mammalian cells: Application to isolation of neutralizing chicken monoclonal anti-IL-12 antibodies. *J. Immunol. Methods* **327**, 40–52 (2007).

58. Zahnd, C., Amstutz, P. & Plückthun, A. Ribosome display: Selecting and evolving proteins in vitro that specifically bind to a target. *Nat. Methods* **4**, 269–279 (2007).

59. Wilson, D. S., Keefe, A. D. & Szostak, J. W. The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3750–3755 (2001).

60. Ghadessy, F. J., Ong, J. L. & Holliger, P. *Directed evolution of polymerase function by compartmentalized self-replication*. www.pnas.orgcgidoi10.1073pnas.071052198.

61. Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci.* **108**, 11399–11404 (2011).

62. Glanville, J. *et al.* Deep sequencing in library selection projects: What insight does it bring? *Curr. Opin. Struct. Biol.* **33**, 146–160 (2015).

63. Wrenbeck, E. E., Faber, M. S. & Whitehead, T. A. Deep sequencing methods for protein engineering and design. *Curr. Opin. Struct. Biol.* **45**, 36–44 (2017).

64. Portnoff, A. D., Stephens, E. A., Varner, J. D. & DeLisa, M. P. Ubiquibodies, synthetic E3 ubiquitin ligases endowed with unnatural substrate specificity for targeted protein silencing. *J. Biol. Chem.* **289**, 7844–7855 (2014).

65. Yau, R. G. *et al.* Assembly and Function of Heterotypic Ubiquitin Chains in Cell-Cycle and Protein Quality Control. *Cell* **171**, 918-933.e20 (2017).

66. Michel, M. A., Swatek, K. N., Hospenthal, M. K. & Komander, D. Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. *Mol. Cell* **68**, 233-246.e5 (2017).

67. Ernst, A. *et al.* A Strategy for Modulation of Enzymes in the Ubiquitin System. *Science (80-.).* **339**, 590–595 (2013).

68. Manczyk, N. *et al.* Structural and functional characterization of a ubiquitin variant engineered for tight and specific binding to an alpha-helical ubiquitin interacting motif. *Protein Sci.* **26**, 1060–1069 (2017).

69. Wyler, E. *et al.* Inhibition of NF-κB activation with designed ankyrin-repeat proteins targeting the ubiquitin-binding/oligomerization domain of NEMO. *Protein Sci.* **16**, 2013–2022 (2007).

70. Grubisha, O. *et al.* DARPin-Assisted Crystallography of the CC2-LZ Domain of NEMO Reveals a Coupling between Dimerization and Ubiquitin Binding. *J. Mol. Biol.* **395**, 89–104 (2010).

CHAPTER 2

REPROGRAMMING A DEUBIQUITINASE INTO A TRANSAMIDASE

This chapter has been published:

Chang, L. H. & Strieter, E. R. Reprogramming a Deubiquitinase into a Transamidase. *ACS Chem. Biol.* **13**, 2808–2818 (2018).[1]

Abstract

Investigating how the different ubiquitin topologies interacting with ubiquitin interacting proteins is crucial to understand the ubiquitin signaling pathway. The main challenge is the accessibility of the ubiquitin conjugates with well-defined ubiquitin chains. Here, we report the engineering of a deubiquitinase, Yuh1 which is capable of auto-ubiquitinating itself site specifically to yield ubiquitin protein conjugate. Through alanine scanning at targeted surfaces followed by random mutagenesis and yeast display-based directed evolution, we have discovered a Yuh1 variant with the transamidation to hydrolysis ratio 28-fold higher than the WT. This enable the robust autoubiquitination of the Yuh1 variant to install distinct ubiquitin chains containing the native isopeptide bond on itself effectively. Further studies using this ubiquitin protein conjugate as the model substrate for deubiquitinases elucidate that certain deubiquitinases activities are affected by the substrates linked to the ubiquitin chains. This emphasizes the importance of interrogation of the ubiquitin protein conjugates in order to fully understand the biological functions of deubiquitinases.

Introduction

Ubiquitination is a post translational modification describing the covalent attachment of ubiquitin moieties onto a substrate through isopeptide or peptide bond. This process is ATP dependent and is catalyzed by an enzymatic cascade involving E1 activating enzymes, E2 conjugating enzymes and E3 ligase enzymes.[2,3] Other than the mono-ubiquitinated substrate, ubiquitin contains amino acids with primary amine, this indicates that ubiquitin can be further ubiquitinated. Together with Met1 and 7 lysines (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63) residues, additional ubiquitin molecules can be attached onto ubiquitin at these primary amines, leading to the generation of polyubiquitin with different linkages. In the circumstances where the same Lys is utilized in the chain elongation, the polyubiquitin is known as homotypic chain; while ubiquitin conjugates containing different linkages are known as heterotypic polyubiquitin. The architectures of these ubiquitin chains are determined by the ubiquitination sites. Studies have shown that these ubiquitin moieties in distinct ubiquitin chains interact differently intramolecularly, hence adopting different conformations.[4,5] This ubiquitination process is reversible and an enzyme, deubiquitinases (DUBs) can remove the ubiquitin modification from the substrates by hydrolyzing the isopeptide bonds.[6,7]

These ubiquitin conjugates regulate a plethora of biological processes in eukaryotes. It has been elucidated that different ubiquitin conjugates trigger distinct biological events. For example, mono-ubiquitinated histone has been associated with transcriptional regulation,[8] Lys48 linked ubiquitinated substrates

are subjected to proteasomal degradation and Lys63 linked ubiquitin protein conjugates are reported to participate in DNA damage response.[9,10] The cellular ubiquitin can exist in the form of unanchored ubiquitin or ubiquitin chains. These free ubiquitin molecules can be generated from (i) the transcription and translation of the ubiquitin gene, (ii) the hydrolyzing product of the ubiquitin conjugates by the DUBs and (iii) the synthesis of unanchored polyubiquitin by certain ubiquitinating enzymes. It has been established that these unanchored polyubiquitins also involve in biological processes. For example, Lys63 linked ubiquitin chains involve in the activation of protein kinases in the NF-κB signaling pathway.[11] Furthermore, quantitative analysis has shown that the percentage of the distinct ubiquitin conjugates vary depends on the cellular states, such as during cell cycles and cellular differentiation.[12]

This project is motivated from our goal to understand how ubiquitin editors and readers recognize these defined ubiquitin conjugates (**Figure 2-1**). To facilitate these studies, enzymatic and chemical methods have been developed to generate ubiquitin activity-based probe, ubiquitin-peptide or protein conjugates (also known as anchored ubiquitin), and unanchored ubiquitin chains. These ubiquitin conjugates have provided valuable insights into the biochemical and biophysical characterizations of the ubiquitin interacting proteins.

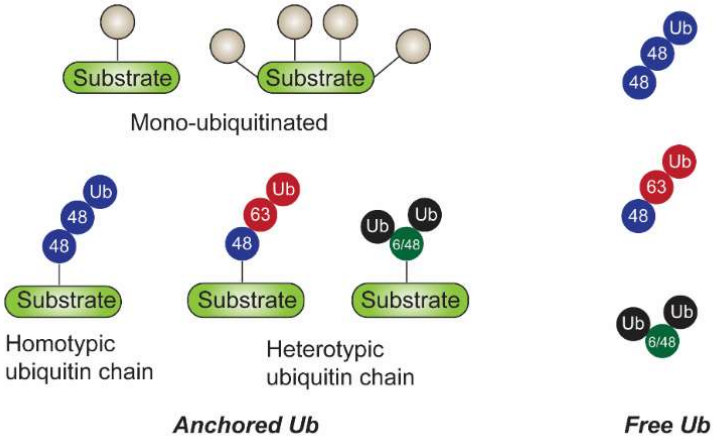

Figure 2-1: The possible topologies of ubiquitin conjugates in cellulo

Despite the advances in the synthesis methods of ubiquitin conjugates, there are limitations with the current techniques. Enzymatic protocol using ubiquitinating enzymes has been widely used to prepare unanchored ubiquitin chains.[13] Nonetheless, the generation of ubiquitin protein conjugates remains challenging and problematic; these ubiquitin protein conjugates tend to be decorated with heterogenous linkages, and the ubiquitination mostly fused to the N or C terminal of the substrate and are non-site specifically.[14,15,16] While the chemical method which modified from native chemical ligation has been successfully applied to prepare ubiquitinated histone site-specifically, the reaction conditions are harsh.[17,18] The necessary denaturing conditions have limited this protocol to be applicable for robust proteins only, such as histone and ubiquitin. Here, we question that could we engineer a modular ubiquitin transamidase to catalyze the transfer of distinct ubiquitin chains on a substrate site specifically to generate ubiquitin protein conjugates of interest?

We have discovered a deubiquitinase, Yuh1, has the potential to be developed as a modular ubiquitin transamidase. Yuh1 is a 26 kDa yeast ubiquitin C-terminal hydrolase which is reported to hydrolyze the isopeptide or peptide bond of ubiquitin conjugate to small adducts. Yuh1 shows structural similarities to papain-like cysteine protease, consisting the catalytic triad at Cys90, His166 and Asp181. Both His166 and Asp181 act as the general base and acid to protonate and prepare Cys90 for the nucleophilic attack at the carbonyl carbon of the scissile bond.[19] It is proposed that upon the binding of Yuh1 with its substrate, for example, ubiquitin with a single Asp extension at its C terminal (UbD77), the Cys90 would attack the scissile bond, leading to the formation of the acyl enzyme intermediate. Under normal conditions, the intermediate will be hydrolyzed to yield the free ubiquitin molecule. Our group has shown that this deacylation step can be intercepted by nucleophilic amines to attach small amines at the ubiquitin C terminal (**Figure 2-2**).[20,21] The transamidation is achieved by extreme reaction conditions; alkaline pH and high concentration of amines are necessary to generate nucleophilic amines to outcompete the prevalent water molecules, and the organic solvent, such as 30% DMSO is supplied as addictive to further reduce the concentration of water to promote the transamidation. Nonetheless, the problem is that Yuh1 still catalyze hydrolysis with an order of magnitude faster than transamidation. The activity of Yuh1 is also impeded in high pH.[22]

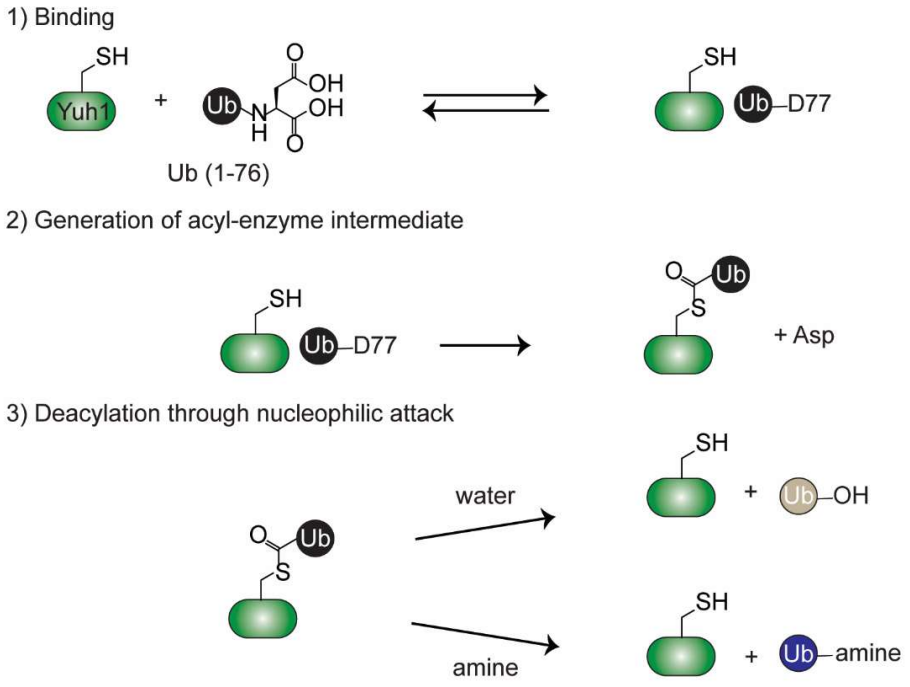

Figure 2-2: The proposed mechanism of Yuh1 catalyzed hydrolysis and transamidation

Here, we decided to reprogram the Yuh1 into a modular ubiquitin transamidase by reducing its hydrolase activity and improving its transamidase activity under milder conditions concurrently. By using an integrated approach of alanine scanning, random mutagenesis and yeast surface display based directed evolution, we have engineered Yuh1 into a ubiquitin transamidase with is capable of auto-ubiquitinate with distinct ubiquitin chains on a single Lys, generating ubiquitin-Yuh1 conjugate with a variety of ubiquitin chains. We further investigated on how different deubiquitinases interact with the ubiquitin-Yuh1 conjugates and elucidated that certain deubiquitinases display markedly different cleavage activities towards ubiquitin-Yuh1 conjugates and unanchored ubiquitin chains. Our result demonstrated that ubiquitin-Yuh1 could serve as a model

substrate to interrogate ubiquitin interacting proteins and suggested that deubiquitinases are sensitive to the conjugated substrates.

Result

Targeted mutagenesis

In previous studies, our group has discovered that a deubiquitinase, Yuh1 is capable to catalyze transamidation to append small amines at the C terminus of mono-ubiquitin or at the proximal ubiquitin of a ubiquitin chains. Nonetheless, this transamidation requires extreme reaction conditions to overcome the hydrolysis. First, high pH and high concentration of the amines are required to generate sufficient nucleophilic amines to outcompete water in the deacylation process. Second, the presence of organic solvent, such as 30% DMSO are necessary to further reduce the concentration of water molecules for the transamidation to occur. The intrinsic hydrolysis activity of Yuh1 is a major hurdle in our Yuh1 engineering effort. Hence, our initial effort is to improve the transamidation to hydrolysis ratio of Yuh1.

We decided to perform targeted mutagenesis at the periphery active site of Yuh1 as we reasoned that these amino acid residues might involve in influencing the catalytic activity of Yuh1. Based on the crystal structure of a ubiquitin conjugated at the catalytic Cys of Yuh1, we have identified 3 regions of interest: (i) the active site- crossover loop (ii) the N terminus of Helix 4 containing catalytic Cys90 and (iii) the β4-β5 loop containing the catalytic Asp181 (**Figure 2-3**). Studies have elucidated that the crossover loop is disordered in the apo Yuh1

but undergoes a major conformational change upon binding with ubiquitin where the loop becomes ordered and contacts with the ubiquitin. In addition, the residues on the N terminus of Helix 4, especially the invariant Asn88, is suggested to stabilize the active site- crossover loop through hydrogen bond. Here, we performed alanine scanning on these regions to investigate the effects of these side chains.

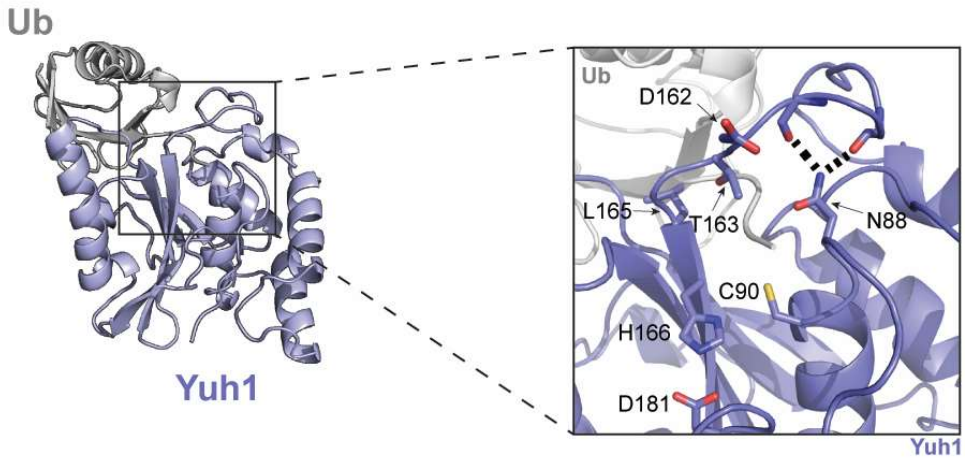

Figure 2-3: Crystal Structure of ubiquitin conjugated at the catalytic Cys of Yuh1 (PDB: 1CMX)

Yuh1 is covalently attached to the ubiquitin aldehyde (Left) with the catalytic triad of Yuh1 (Cys90, His166 and Asp181) highlighted (Right).

We expressed the Yuh1 variants in *E.coli* and we performed the transamidation reaction using the crude purified Yuh1 variants with UbD77 as substrate and allylamine as nucleophile. The reactions were carried out under milder conditions, at pH 8 (contrast to pH 10.5) and for an extended time (**Figure 4A**). To rapid screen for the transamidation and hydrolysis yield of the Yuh1 variants, we have adopted a matrix-assisted laser desorption/ionization- time of

flight mass spectrometry (MALDI-TOF MS) assay which was designed originally to quantify *in vitro* DUB activity.[23] In short, the amount of hydrolysis and transamidation products were quantitated using ^{15}N-labeled UbD77 as the internal standard (**Figure 2-4B**). Calibration curves for both hydrolysis product, UbWT and transamidation product, Ub allylamine (UbAA) confirmed the proportionality of the assay, with the detection capability spanning a range of 10^2-10^4 nM (**Figure 2-5A and 5B**).

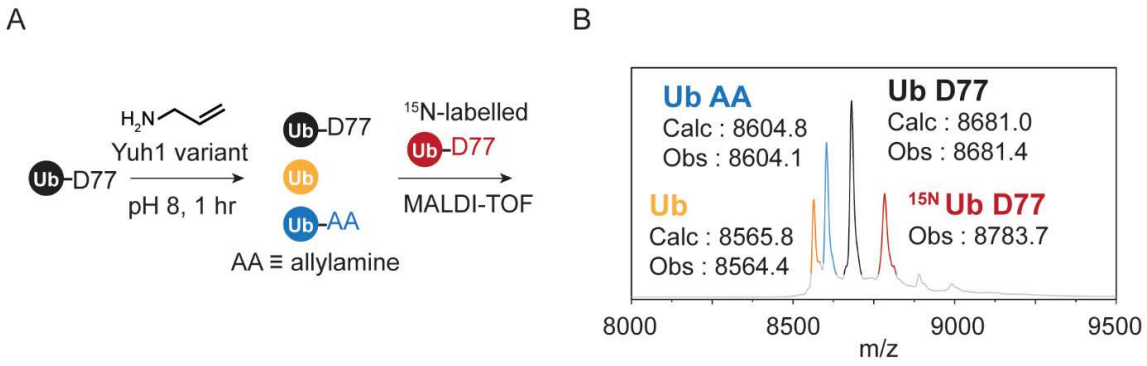

Figure 2-4: MALDI-TOF MS assay

(**A**) Reaction conditions and the sample preparation for the MALDI-TOF assay to screen for the transamidation and hydrolysis yield for the Yuh1 variants. (**B**) An example of acquired MALDI-TOF MS spectrum. Products were quantified using the spiked-in ^{15}N-labeled UbD77 as the internal standard.

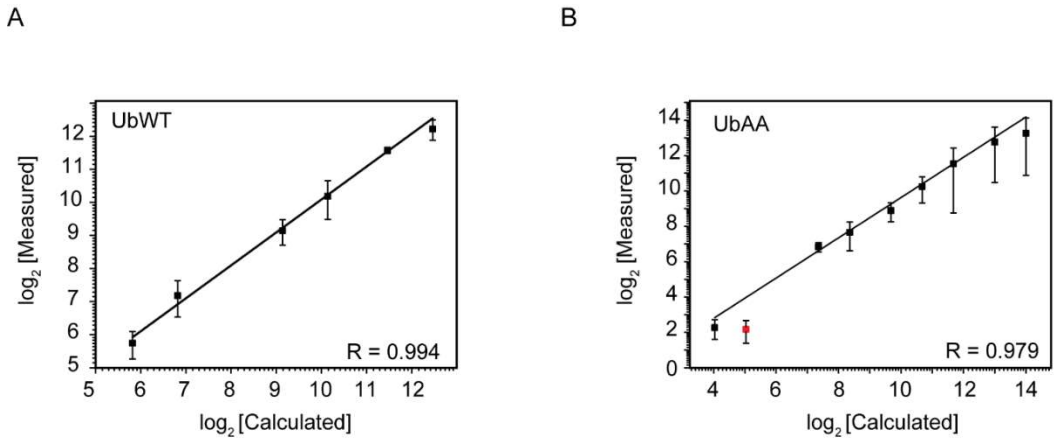

Figure 2-5: Calibration curve for MALDI-TOF MS assay

Calibration plot for (**A**) Ub WT and (**B**) Ub-allylamine using MALDI-TOF MS. Data points were obtained from the average measured concentration of 2 technical replicates from 2 biological replicates. Fitting was performed in OriginPro and linearity was conserved over 2 orders of magnitude. Red dot represented outlier.

Our result showed that under these reaction conditions, wild type Yuh1 (Yuh1 WT) did not yield any transamidation product, presumably the generated UbAA was rapidly hydrolyzed into UbWT. However, with a single N88A mutation, this Yuh1 variant starts to display higher transamidase activity. Among the alanine mutants, we found that N88A/D162A/L165A (herein referred as Yuh1m) exhibits the most positive improvement, with a transamidation yield around 30% and a transamidation to hydrolysis ratio at 3. Both Asp162 and Leu165 are located near the C terminus of the active-site crossover loop (**Figure 2-3**). While alanine substitution of the residues at the β4-β5 have negative effect on the transamidation activity of Yuh1 N88A, as the transamidation to hydrolysis ratio has decreased to less than 0.5 compared to Yuh1 N88A. Thus, our result indicated that alanine mutation of Asn88 and crossover loop significantly improve the transamidation activity of Yuh1 (**Figure 2-6A and 6B**).

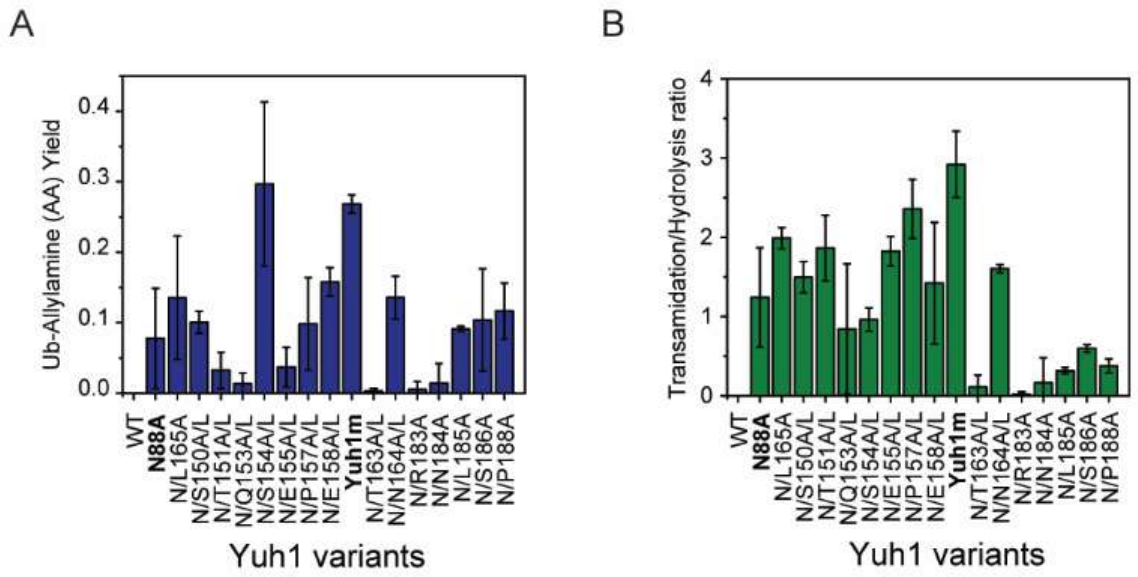

Figure 2-6: Transamidation of Yuh1 variants.

(**A**) Yield of Ub-allylamine (UbAA) and (**B**) transamidation to hydrolysis ratio of Yuh1 variants determined through MALDI-TOF MS. Error bars correspond to standard deviation of two biological replicates and two technical replicates.

To gain insight into the improved transamidation activity of these Yuh1 variants, we performed initial rate measurements, focusing at Yuh1 WT (without and with the presence of amine), Yuh1 N88A, and Yuh1m. The Michaelis-Menten kinetics showed that for Yuh1 WT, the catalytic efficiency for hydrolysis under both circumstances are similar, with 6±3 x 10^5 $M^{-1}s^{-1}$ without amine and 8±2 x 10^5 $M^{-1}s^{-1}$ with amine, with both *K*m and k_{cat} are at the same order. This suggested that the catalytic efficiency for hydrolysis is not affected by the presence of amines. While the catalytic efficiency for transamidation is 9±1 x10^4 $M^{-1}s^{-1}$, which is 10 times lower compare to hydrolysis. The low catalytic efficiency is mainly due to the decrease in K_m (**Figure 2-7A and Table 2-1**).

With the introduction of N88A, we observed a significant decrease in the rate of catalytic turnover and catalytic efficiencies for both hydrolysis and transamidation. Especially for hydrolysis, the k_{cat} and catalytic efficiency for Yuh1 N88A is 0.062±0.006 s^{-1} and 2.3±0.6 x 10^3 $M^{-1}s^{-1}$, which are close to two order of magnitude lower compare to Yuh1 WT. For transamidation, it is close to 10 times lower compare to Yuh1 WT, at 3±1 x 10^3 $M^{-1}s^{-1}$. As a result, the catalytic efficiency for transamidation is comparable to the catalytic efficiency for hydrolysis (**Figure 7B and Table 1**). We detected the similar trend with Yuh1m, with the similar overall decrease in the rate of catalytic turnover and catalytic efficiencies, but with the catalytic efficiency for transamidation to hydrolysis at a ratio of 1.2. Especially the catalytic efficiency of transamidation for Yuh1m is at 6±2 x 10^3 $M^{-1}s^{-1}$, which is slightly higher compare to N88A (**Figure 7C and Table 1**). Together, these data demonstrate that these alanine mutations have compromised the catalytic activity of Yuh1, renders it an ineffective and yet promiscuous enzyme, making the Yuh1 N88A and Yuh1m a better transamidase compare to Yuh1 WT. This is not a unique phenomenon, as promiscuity has reported to be a key factor in the evolution of new protein function.[24,25]

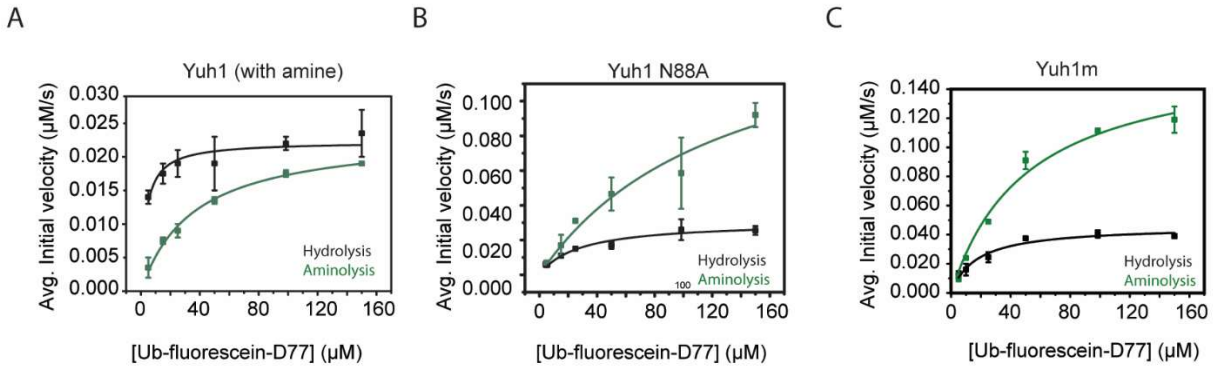

Figure 2-7: Michaelis-Menten kinetics measurement of Yuh1 variants.

Steady-state kinetics analysis of transamidation and hydrolysis for (**A**) Yuh1 WT, (**B**) Yuh1 N88A and (**C**) Yuh1m in the presence of allylamine. Error bars represent the standard error of two biological replicates.

Table 2-1: Michalis-Menten kinetics Parameters for hydrolysis and transamidation with Yuh1 variants.

	Hydrolysis			**Transamidation**		
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
Yuh1 (no amine)	5±2	3.1±0.2	6±3 X 10^5	-	-	-
Yuh1	4±1	3.0±0.1	8±2 X 10^5	36±4	3.1±0.1	9±1 X 10^4
Yuh1 N88A	27±7	0.062±0.006	2.3±0.6 X 10^3	130±60	0.32±0.08	3±1 X 10^3
Yuh1m	17±4	0.092±0.006	5±1 X 10^3	51±11	0.34±0.02	6±2 X 10^3

Error represent the standard error of two biological replicates.

Autoubiquitination of Yuh1m

One of the properties of the ubiquitinating enzymes is that they are capable of auto-ubiquitination.[26] Here, having discovered a Yuh1 variant with improved transamidation and impaired hydrolysis, we would like to investigate whether Yuh1m is capable to self-ubiquitinate if its Lys is at close proximity. We

decided to mimic the effect by introducing a lysine at the active site- crossover loop which is at periphery of the Yuh1 active site. We chose Thr163, as the crystal structure suggested that its side chain is pointing towards the catalytic Cys, and is within 4 Å from the C terminus of ubiquitin (**Figure 2-3**).

We have generated the following Yuh1 variants: Yuh1 T163K and Yuh1m T163K (herein referred as Yuh1qm) and reacted the Yuh1 variants with high concentration of UbD77. SDS-PAGE analysis showed that only for Yuh1qm, there are appearance of multiple protein bands at higher molecular weight. This is accompanied with the decrease in the band representing Yuh1qm. The most intense band (Band 1) has a molecular weight corresponding to Yuh1qm modified with a single ubiquitin. While Band 2 has a higher molecular weight compare to Band 1, its molecular weight is lower than a di-ubiquitinated Yuh1qm; Band 3 with the highest molecular weight is consistent with a Yuh1qm decorated with two ubiquitins (**Figure 2-8A**). We further validated these bands represent the auto-ubiquitinated Yuh1qm through Western blot analysis (**Figure 2-8B**). On the other hand, we observed no auto-ubiquitinating species for Yuh1 WT, Yuh1 T163K and Yuh1m. Next, we would like to investigate whether Yuh1qm can auto-ubiquitinate with pre-assembled ubiquitin chains. Here, we reacted Yuh1qm with K48- and K63-linked ubiquitin dimers with UbD77 as the proximal ubiquitin. As predicted, Yuh1qm rapidly auto-ubiquitinate with the ubiquitin dimers regardless of the linkages (**Figure 2-8C**). Contrast with the previous auto-ubiquitination assay with mono-ubiquitin, we only observed a single band corresponds to the auto-ubiquitinating species, suggesting that Yuh1qm could not multi-ubiquitinate

using ubiquitin chain as substrate. Here, we acknowledged that we could not discriminate the fact that Yuh1qm could ubiquitinate another Yuh1qm, but the inter-ubiquitination is a entropically unfavorable process.[27]

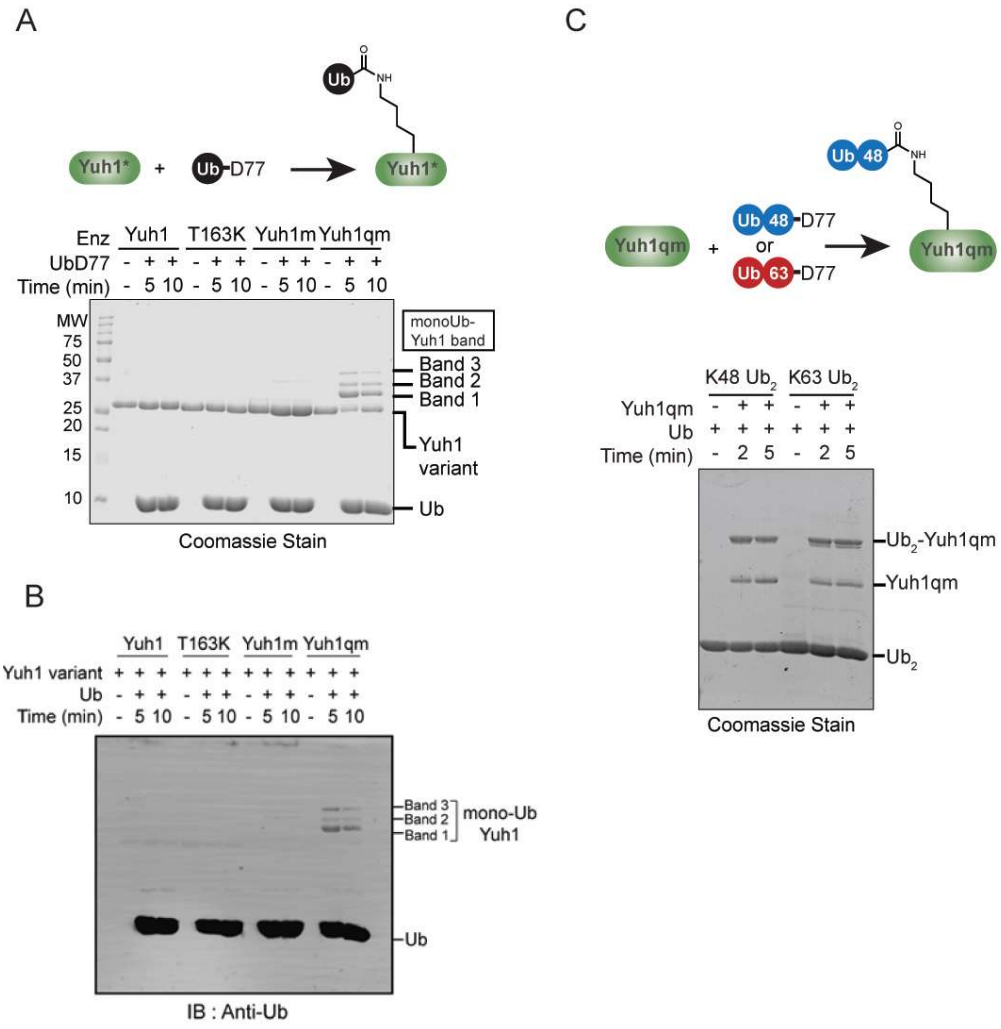

Figure 2-8: Autoubiquitination of Yuh1 variants

SDS-PAGE analysis of auto-ubiquitination of Yuh1 variants (50μM) with UbD77 (200 μM) using (**A**) Coomassie blue and (**B**) Western blot with anti-Ub. The new protein bands with higher molecular weight band represent auto-ubiquitinating Yuh1qm. (**C**) SDS-PAGE analysis of auto-ubiquitination of Yuh1qm (5 μM) with K48- and K63-linked ubiquitin dimers (20 μM).

To identify the ubiquitination site, we performed tandem mass spectrometry (MS/MS). As result indicated that Yuh1qm is capable of multi-monoubiquitinated when UbD77 is the substrate, we decided to perform an in-gel digestion on bands containing auto-ubiquitinating species with a combination of trypsin and GluC (**Figure 2-8A**). We detected two peptides containing the diGly-modified (which is the signature of ubiquitination) Lys. They are peptides containing the expected T163K and K66 which is located on a poorly resolved region in the Yuh1 crystal and at the close proximity to the active site of Yuh1 (**Figure 2-9A**). While the MS/MS analysis for the purified di-ubiquitinated Yuh1qm showed that K163 is the only Lys that is ubiquitinated (**Figure 2-9B**).

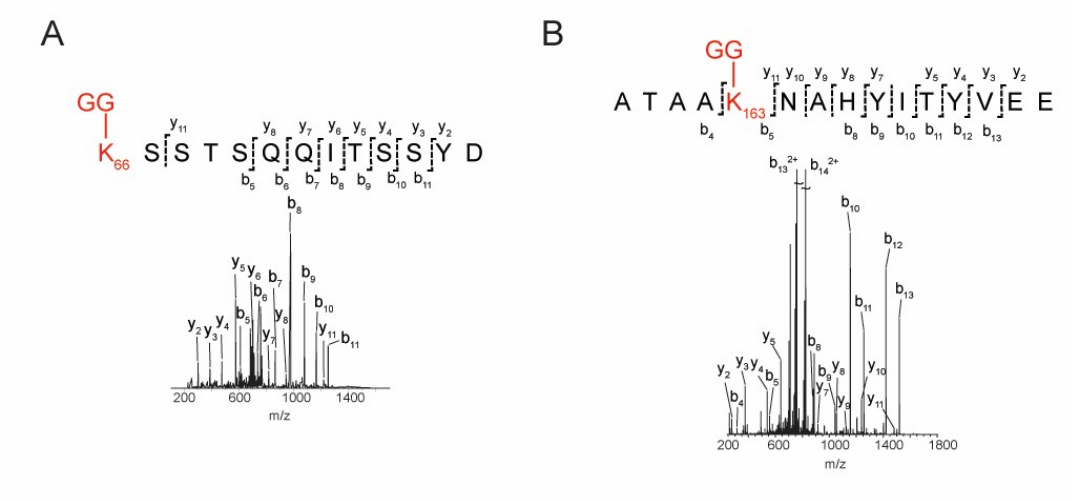

Figure 2-9: Tandem mass spectrometry analysis of auto-ubiquitinating Yuh1qm

MS/MS spectrum showing peptides with ubiquitination site at (**A**) K66 and (**B**) T163K of Yuh1qm.

Our result has elucidated that Yuh1qm can auto-ubiquitinate site specifically regardless of the ubiquitin topology. Hence, Yuh1qm has the potential to be developed as the modular ubiquitin transamidase to generate ubiquitin-protein conjugate, that is the ubiquitin-yuh1 conjugate. However, a time course assay using lower concentration of ubiquitin, which is 100 times lower compare to the initial assay, revealed that hydrolysis of the auto-ubiquitinating Yuh1qm occur with extended reaction times (**Figure 2-10A and Figure B-1A**). The yield of the auto-ubiquitinating species is around 30% at the initial time point, and the measured half-life is estimated to be 160 ± 60 s (**Figure 2-10B**). As a result, large scale synthesis of ubiquitin-yuh1 conjugate with Yuh1qm remains challenging.

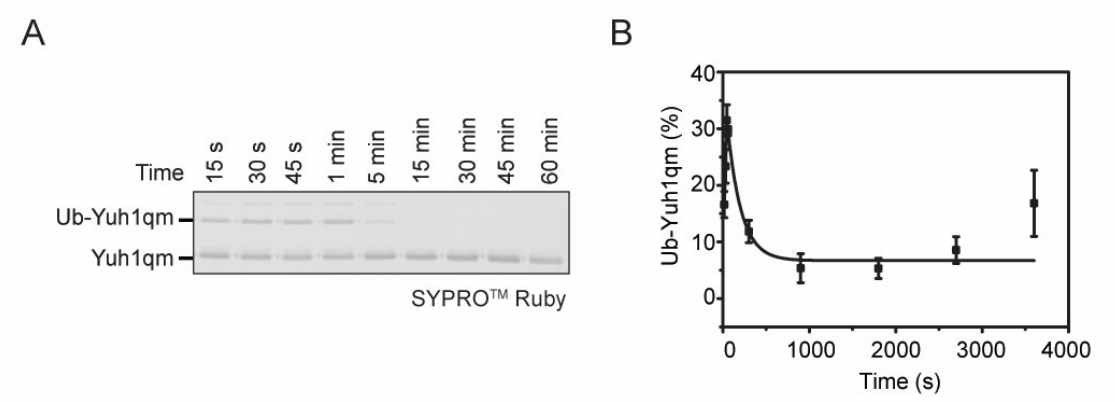

Figure 2-10: Time course assay for auto-ubiquitinating Yuh1qm

(**A**) SDS-PAGE analysis of the auto-ubiquitinating Yuh1qm (1 μM) and UbD77 (2 μM) using SYPRO™ Ruby. (**B**) Scatter plot of the auto-ubiquitinating species. Hydrolysis data was fit to a single exponential decay. Error bars represent the standard error of three biological replicates.

Directed evolution of Yuh1qm

The poor stability of the auto-ubiquitinating Yuh1qm has precluded us from applying Yuh1qm to synthesis ubiquitin-yuh1 conjugate in large scale. Here, we decided to perform further engineering on Yuh1qm through yeast surface display based directed evolution. An ideal Yuh1qm variant would have improved auto-ubiquitination efficiency and further reduced intrinsic hydrolytic property.

First, we would like to investigate whether yeast surface display serves as an ideal platform for our directed evolution strategy. In our strategy, we chose to display Yuh1qm variants on the yeast surface through the well-established yeast Aga1p-Aga2p display system, with the displayed Yuh1 variant flanking between a HA and a c-Myc tag, which could be detected using anti-HA or anti-cMyc.[28,29,30] The Yuh1qm variants were then treated with biotinylated UbD77. We hypothesized that the Yuh1qm variants with desired properties would have higher biotin labeling on the cell surface and could be detected through streptavidin. Under the circumstances where the fluorescent labeled antibodies and streptavidin were used, the labeled yeast cell could be screened by fluorescent-activated cell sorting (FACS) to identify and sort the most fluorescent yeast cells (**Figure 2-11A**). To validate our strategy, we first transformed and induced the yeast cells with yeast display vector containing either Yuh1 WT or Yuh1qm. The cells were incubated with biotinylated UbD77 and and flow cytometry analysis showed that only Yuh1qm is capable to label itself with biotinylated ubiquitin (**Figure 2-11B**). Here, we would like to mention that labeling at cMyc tag using antibody has negatively impacted the degree of auto-

ubiquitination (aata not shown). Hence, although labeling at the C terminus cMyc tag would ensure the identification of non-truncated Yuh1qm variants, we decided to label the HA tag at the N terminus of Yuh1qm variants to estimate the degree of Yuh1qm expression on the yeast surface.

We have generated a random mutagenized Yuh1qm library through error-prone PCR (epYuh1qm) with an estimated size of 10^7. Sanger sequencing of 10 clones showed that the library members contain 1-3 mutations per gene, with mutated residues distributed evenly across Yuh1qm. We have performed five rounds of screenings; in each screening, we were isolating the clones with both high expression level of Yuh1 and high degree of ubiquitination. The sorted clones were amplified and subjected for the subsequent rounds of sorting. In addition, with each round we increased the stringency of the reaction conditions by lowering the concentration of biotinylated UbD77 (**Figure 2-11C**). In addition, the sorted yeast cells were incubated with biotinylated ubiquitin and the degree of binding were investigated using flow cytometry. This was to ensure we did not enrich for clones with higher binding affinities towards ubiquitin (data not shown).

At round 3, we started to observe the appearance of populations with higher degree of ubiquitination. Sanger sequencing revealed that these clones containing a consensus mutation at N164K. N164K is located at the active site-crossover loop adjacent to the previously introduced T163K. Further sorting showed the additional of two mutation hotspots; one at the periphery to the catalytic Cys90 (Val86 and Leu139) and one is distant to both the active site and the binding interaction surfaces with ubiquitin (Val52 and Ile202) (**Figure 2-12**).

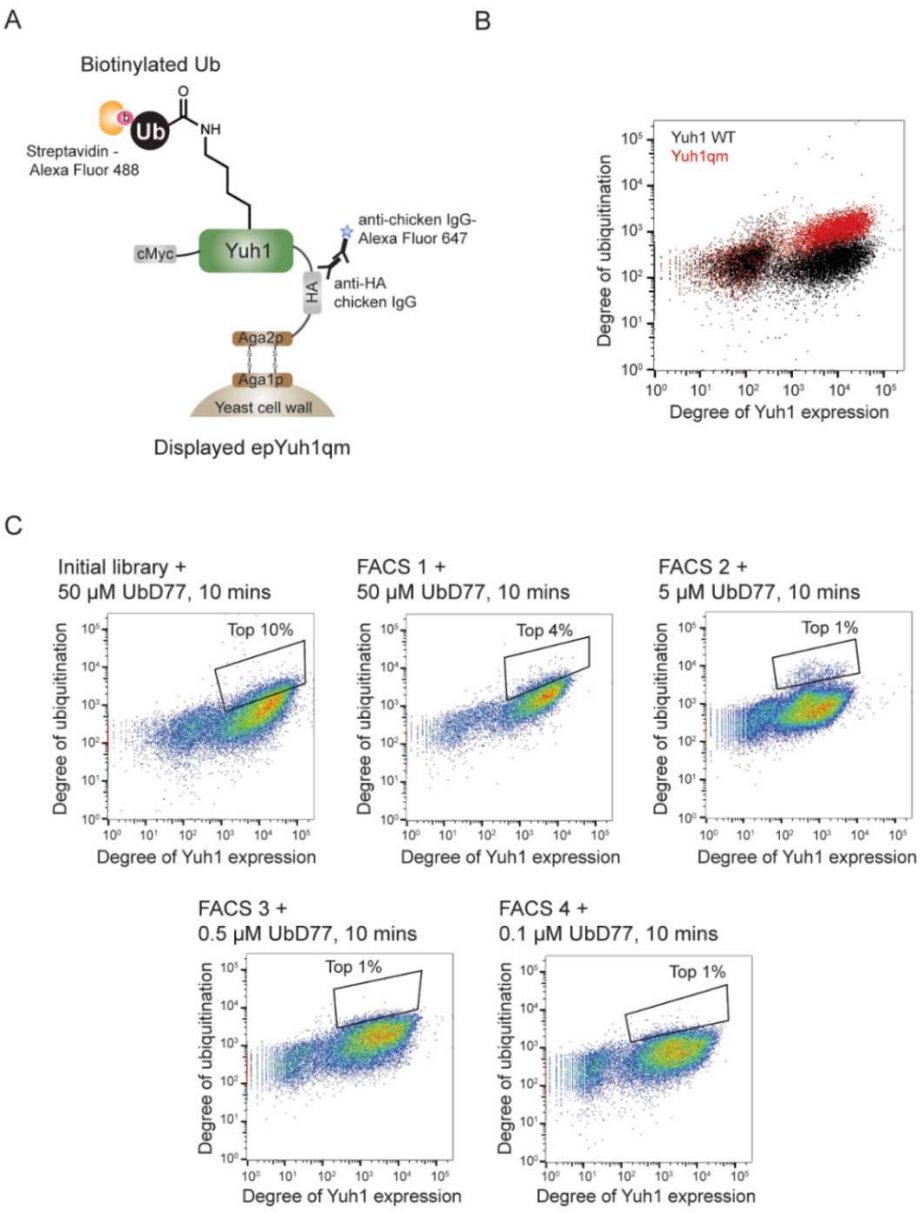

Figure 2-11: Directed evolution of Yuh1qm

(**A**) Scheme of yeast surface display for the directed evolution of Yuh1qm. The displayed Yuh1 variants were detected with the labeling of the HA-tag fused to the N terminus of Yuh1 with chicken anti-HA and Goat anti-chicken conjugated to AlexaFluor® 647. The sorting of the library members was based on the degree of ubiquitination on the Yuh1 which was immobilized on the yeast surface. The ubiquitin was biotinylated and could be detected using streptavidin conjugated to Alexa Fluor® 488. (**B**) Flow cytometry dot plot of yeast cells expressing Yuh1 WT and Yuh1qm incubated with biotinylated UbD77. (**C**) FACS of the random mutagenized epYuh1qm library. The frames and percentages designate the fraction of yeast population isolated in each round of screening.

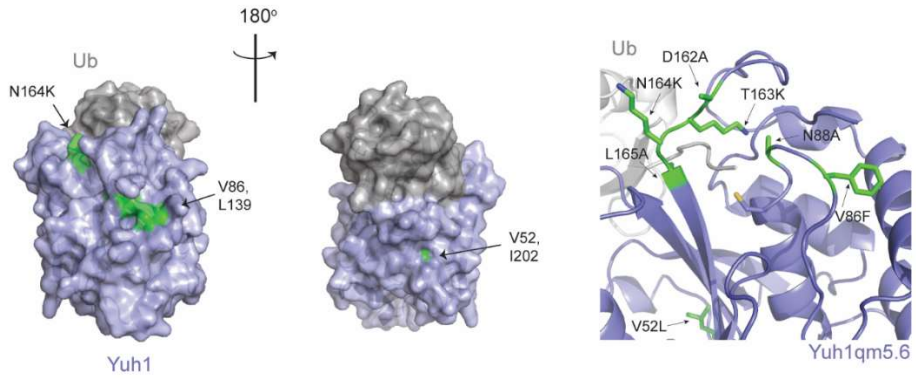

Figure 2-12: Mutational hotspots

Crystal structure of Yuh1 (PDB ID: 1CMX) with mutational hotspots showing in green (Left) and with the mutated residues highlighted (Right).

Based on the sequencing result, we further generated and expressed these subsequent clones: Yuh1qm N164K (referred as Yuh1qm3.1), Yuh1qm3.1 V86F (Yuh1qm5.1), Yuh1qm3.1 V52L (Yuh1qm5.3) and Yuh1qm V52L/V86F (Yuh1qm5.6). We investigated the auto-ubiquitination efficiency of these clones through the aforementioned time course assay. Consistent with our screening result, all the auto-ubiquitinating species for the engineered Yuh1qm variants are significantly more robust compare to Yuh1qm (**Figure 2-13A and Figure B-1**). Especially for Yuh1qm5.6, the yield for auto-ubiquitinating species approaching 90% at the initial time point (**Figure 2-13B**). Despite the improvement in the auto-ubiquitination efficiency, the engineered Yuh1 variants still experienced auto-hydrolysis. Nonetheless, the engineered Yuh1qm variants are more resistant to hydrolysis which remove the ubiquitin modification from itself. Our result indicated that the yield of auto-ubiquitinating species for Yuh1qm3.1, Yuh1qm5.1, and Yuh1qm5.6 are around 30% even after 1 hr, while for the Yuh1qm, it could

be barely detected (**Figure 2-13A and Appendix Figure B-2**). Fitting to single exponential decay curves estimated the half-lives for auto-ubiquitinate Yuh1qm5.1, Yuh1qm5.3, and Yuh1qm5.6 are 1400±100 s, 500±100 s and 800±100 s (**Figure 2-13B**). We could not fit the single exponential decay curve for the auto-ubiquitinate Yuh1qm3.1.

The overwhelming enrichment of Yuh1qm variants containing the consensus mutation at N164K leaded us to question the role of N164K in stabilizing the auto-ubiquitinating species. Crystal structure of Yuh1 suggesting that the carboxamide side chain of the Asn164 is pointing away from the catalytic Cys90 (**Figure 2-1A**). We surmised that the N164K might be the preference auto-ubiquitination site; upon ubiquitination, the C terminus of ubiquitin would distance itself from the catalytic Cys90, hence reduce the degree of auto-hydrolysis. Indeed, a MS/MS analysis of purified auto-ubiquitinated Yuh1qm5.6 elucidated that N164K is the only ubiquitination site, supporting our supposition (**Figure 2-13C**).

We further performed a Michaelis-Menten kinetics of Yuh1qm V52L/V86F with UbD77 as substrate and allylamine as the nucleophile to investigate the effect of replacing the Val residues with more hydrophobic and bulkier Leu or Phe residues. The data analysis elucidated that while the steady-state kinetics parameters for hydrolysis remain constant, for transamidadtion, the k_{cat} reduced by 45%, the K_m decreased by 4-fold. Together, the catalytic efficiency for transamidation improve close to 3 times compare to Yuh1m, and with a catalytic efficiency transamidation to hydrolysis ratio at 3 (**Figure 2-14 and Table 2-2**).

One of the possible reasons is that the higher degree of hydrophobicity at the periphery of the active site enhances the interaction with the nucleophilic amine at that region, thereby improves the transamidation efficiency.[31]

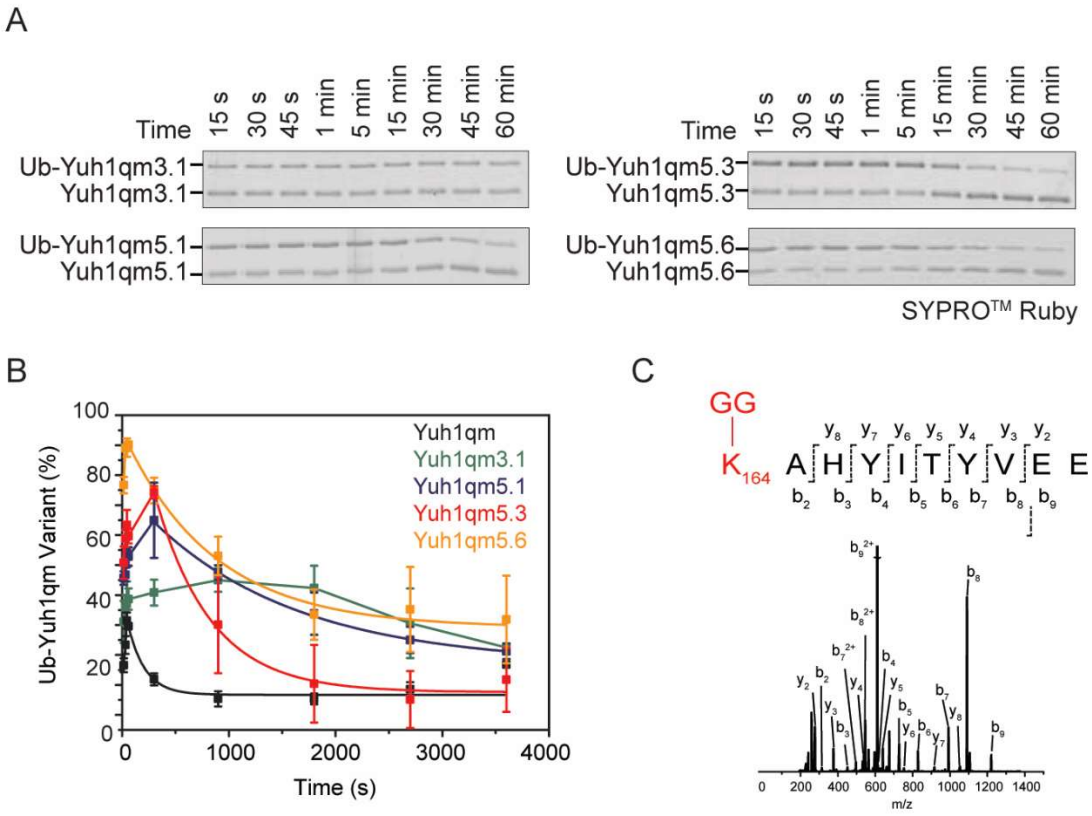

Figure 2-13: Characterization of Yuh1qm variants.

(**A**) SDS-PAGE analysis of the auto-ubiquitinating Yuh1qm variants (1 μM) and UbD77 (2 μM) using SYPRO™ Ruby. (**B**) Scatter plot of the auto-ubiquitinating species. Hydrolysis data was fit to a single exponential decay. Error bars represent the standard error of three biological replicates. (**C**) MS/MS spectrum showing peptides with ubiquitination site at N164K of Yuh1qm5.6.

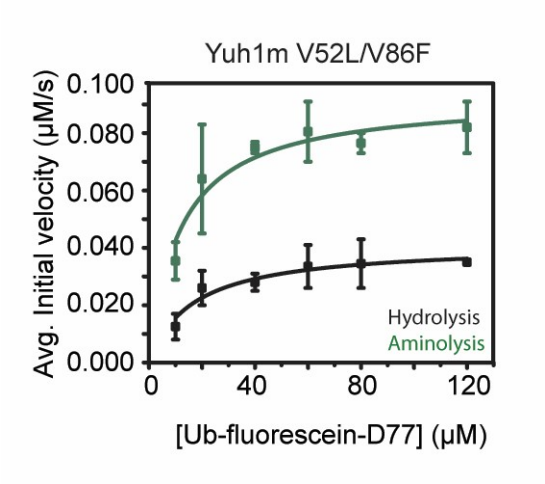

Figure 2-14: Michaelis-Menten kinetics measurement of Yuh1m V52L/V86F.

Steady-state kinetics analysis for transamidation and hydrolysis in the presence of allylamine. Error bars represent the standard error of two biological replicates.

Table 2-2: Michalis-Menten kinetics Parameters for hydrolysis and transamidation with Yuh1m V52L/V86F.

	Hydrolysis			**Transamidation**		
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (x10^3 $M^{-1}s^{-1}$)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (x10^3 $M^{-1}s^{-1}$)
Yuh1m V52L/V86F	16±4	0.082±0.006	5±1	12±3	0.19±0.02	16±2

Error represent the standard error of two biological replicates.

Profiling DUBs activities using Ubiquitin-Yuh1 conjugates

We have shown the discovery of a Yuh1 variants, Yuh1qm5.6 that is capable to auto-ubiquitinate with distinct pre-assembled ubiquitin chains site specifically with high yield. Here, we generated ubiquitin-yuh1 conjugates with different linkages as a model ubiquitin-protein conjugate substrate to study their interactions with ubiquitin editors and readers. To ensure the ubiquitin-yuh1 conjugates do not auto-hydrolyze in the downstream biochemical assay, these

ubiquitin-protein conjugates were treated with *N*-ethylmaleimide (NEM) to quench the Yuh1qm5.6 after auto-ubiquitination.

Here, we investigated how the ubiquitin chain anchoring affects the activities of deubiquitinases. As a proof of study, we synthesized free K48 ubiquitin dimers (Ub_2) and K48 Ub_2-Yuh1 conjugates, or known as the anchored K48 Ub_2. Both free and anchored K48 Ub_2 were reacted with USP11, USP15 and USP21.[32,33,34] These deubiquitinases are members of the ubiquitin specific protease (USP) family and are non-linkage specific. The chain cleavage activities were determined through a gel-based assay. For anchored K48 Ub_2, Western blotting with anti-His was performed to quantitate the disappearance of the K48 Ub_2-His_6-Yuh1qm5.6 and the appearance of mono-ubiquitin-His_6-Yuh1qm5.6 and His_6-Yuh1qm5.6 simultaneously. For free K48 Ub_2, SYPRO™ Ruby stain was performed to determine the disappearance of K48 Ub_2 and the appearance of mono-ubiquitin. Our result elucidated that all USPs cleave the free and anchored K48 Ub_2 with similar efficiencies, although it was suggested that USP11 and USP15 processed free K48 Ub_2 at higher rate. Interestingly, all the USPs could not remove the proximal ubiquitin from the Yuh1qm5.6 (**Figure 2-15A and Figure B-2**). To ensure that NEM treatment did not interfere with the activities of USP, we reacted the non-NEM treated anchored K48 Ub_2 with USP15. Our result suggested that the rate of hydrolysis was approximately the same (**Figure B-3A**).

In addition, we also prepared free and anchored K63 Ub_2 and performed the same DUB assays. Our result was consistent to the data obtained using the free and anchored K48 Ub_2 (**Figure 2-15B**). These observations indicated that

there is specificity towards the conjugated substrate, and implied that USP could process the anchored ubiquitin chain from the distal end regardless of the substrate.

Next, we profile the OTUB1 and OTUB2 activities with the free and anchored Ub_2. Both deubiquitinases belong to the ovarian tumor (OTU) families, which are characterized by their linkage specific activities. OTUB1 specifically processes K48-linked ubiquitin chain while OTUB2 selectively cleaves K48- and K63-linked polyubiquitin.[35,36,37] The activity of OTUB1 is reported to be enhanced by the presence of E2 conjugating enzyme and an OTUB1-UBE2D2 fusion protein was used in the DUB assay.[38] As predicted, the free Ub_2 were rapidly processed by the OTU enzymes (**Figure 2-16A and 16B**). Surprisingly, the OTU enzymes, even the improved OTUB1 variant could not hydrolyze the anchored K48 Ub_2 (**Figure 2-16C**). This was not due to the inhibition of Yuh1qm5.6, as the free K48 Ub_2 was hydrolyzed by the OTU enzymes in the presence of anchored K48 Ub_2 (**Figure B-3B**). On the other hand, OTUB2 is still capable to cleave the anchored K63 Ub_2 (**Figure 2-16D**).

Majority of the K48 linked Ub_2 population adopt a 'compact' conformation; this might lead to the difficulty of the OTU enzymes in accessing the anchored ubiquitin subunits and hence, the lack of the ubiquitin chain cleavage activities. Thus, we extended the chain length and the branching level of the anchored polyubiquitin. We synthesized pre-assembled K48 Ub_3 and K6/K48 Ub_3 and transferred them onto Yuh1qm5.6 and profiled OTUB1 and OTUB2 activities towards these anchored ubiquitin chain.[39] Again, we observed no cleavage

activity towards the anchored ubiquitin chain, but not the free counterpart (**Figure 2-17A and 17B**). Our data indicated that OTUB1 and OTUB2 activities are influenced by the chain anchoring effect. Crystal structure of the OTUB1 conjugated to a ubiquitin aldehyde and UBE2D2~ubiquitin thioester intermediate has been reported to mimic the interaction of OTUB1 with K48 Ub_2. A closer look at the crystal structure suggested that substrate conjugated to K48-linked ubiquitin chains might sterically hinder the OTUB1, thereby preventing the anchored ubiquitin chain processing by OTUB1.[40]

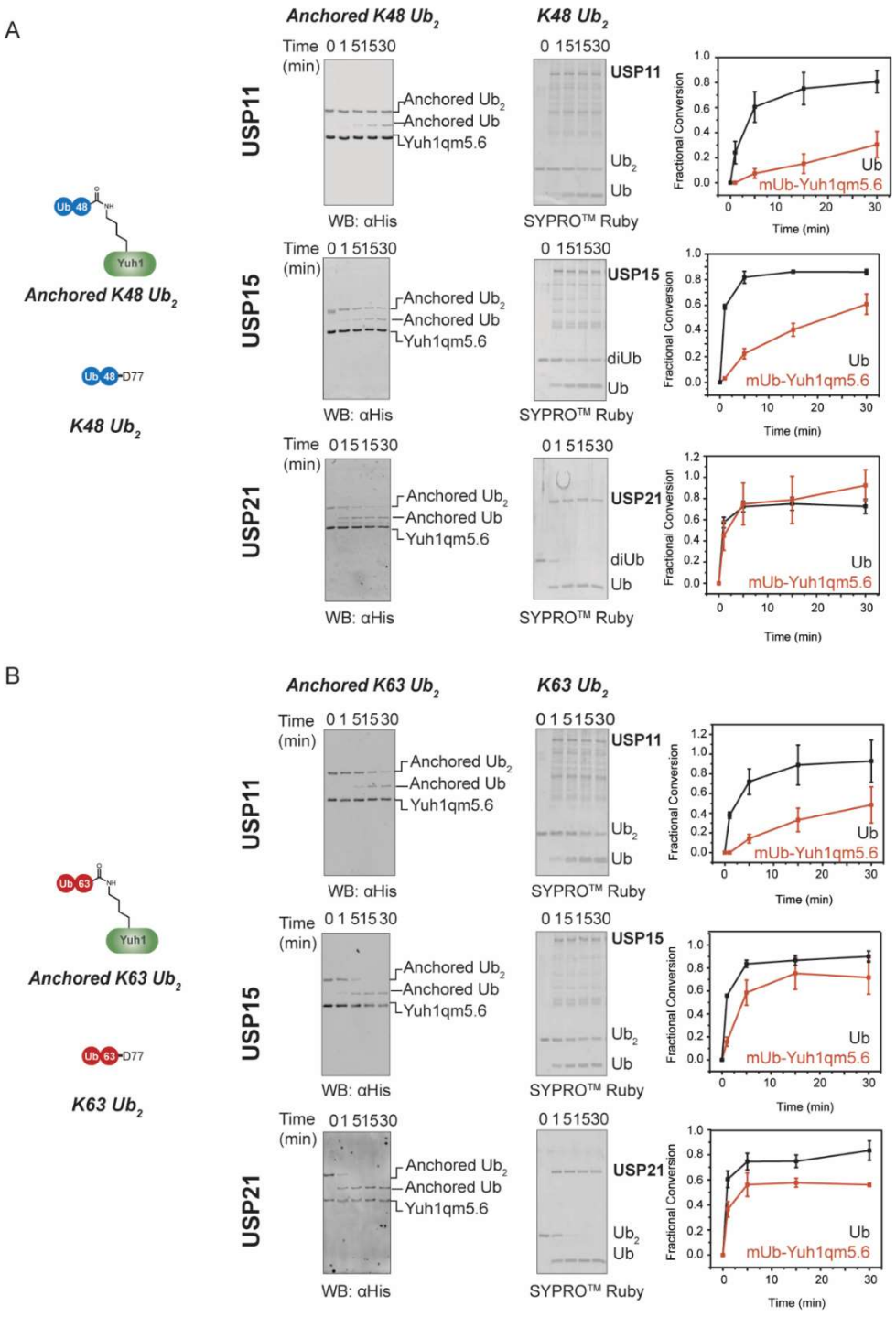

Figure 2-15: Profiling USP activities against free and anchored ubiquitin dimers.

Cleavage of free and anchored (**A**) K48 Ub_2 and (**B**) K63 Ub_2 by USP11, USP15 and USP21. Error bars corresponded to standard error of three biological replicates.

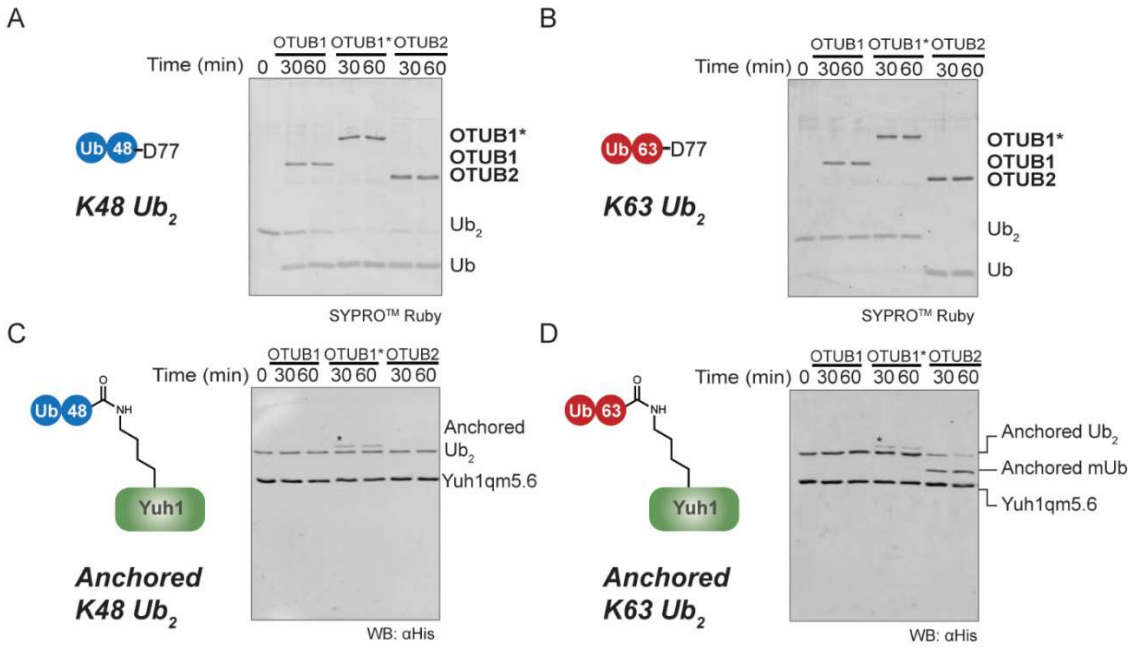

Figure 2-16: Profiling OTU activities against free and anchored ubiquitin dimers.

Cleavage of free (**A**) K48-linked and (**B**) K63-linked ubiquitin dimers and anchored (**C**) K48-linked and (**D**) K63-linked ubiquitin dimers by OTUB1, OTUB1-UBE2D2 (labeled as OTUB1*) and OTUD2.

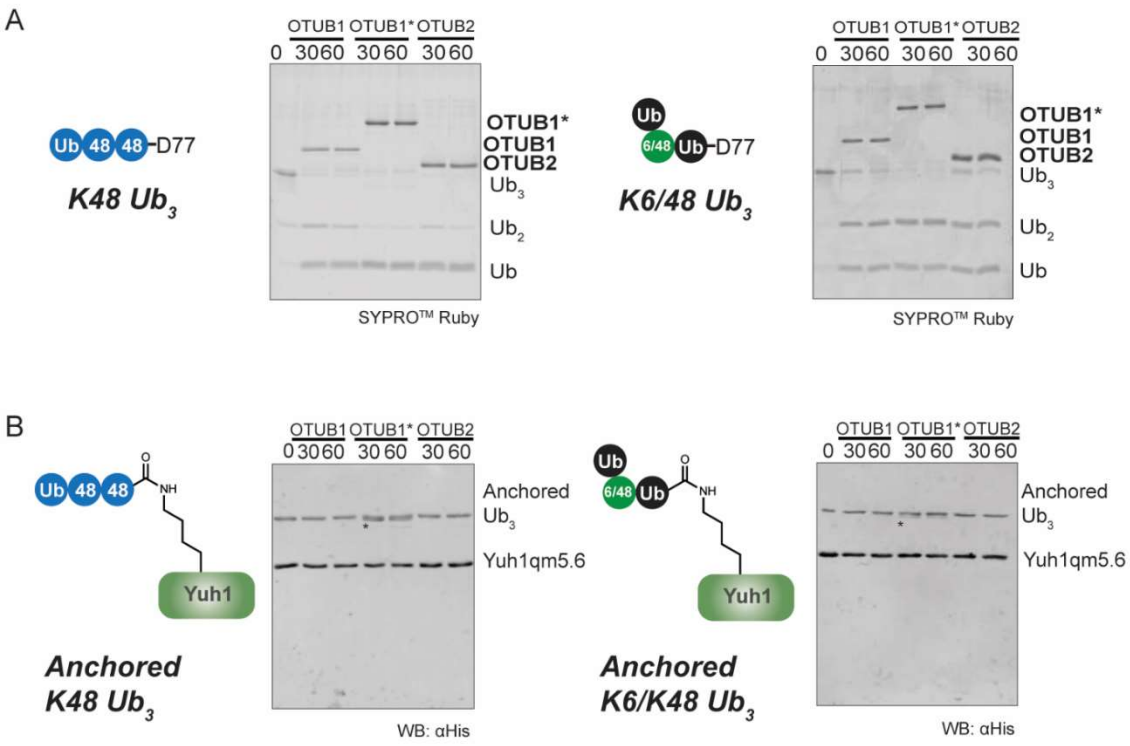

Figure 2-17: Profiling OTU activities against free and anchored ubiquitin trimers.

Cleavage of free (**A**) K48-linked and (**B**) K6/48-linked ubiquitin trimers and anchored (**C**) K48-linked and (**D**) K6/48-linked ubiquitin trimers by OTUB1, OTUB1-UBE2D2 (labeled as OTUB1*) and OTUD2.

Discussion and Future Directions

In this study, we have reprogrammed a deubiquitinase, Yuh1 into a ubiquitin transamidase which is capable to install defined ubiquitin chains on itself site specifically. We would like to emphasize that the ubiquitinating mechanism catalyzed by the Yuh1 is different to the conventional ATP dependent enzymatic cascade, which require the generation of a high energy ubiquitin~AMP, followed by transthioesterification between the catalytic Cys of ubiquitinating enzymes before discharged onto the Lys of the substrate of interest.[41] For our engineered Yuh1, the activated ubiquitin species is resulted

from the nucleophilic attack of ubiquitin with small adduct at the C terminal with a highly reactive catalytic Cys. Hence, the application of reprogrammed Yuh1 in preparing ubiquitin protein conjugate which bypassing the requirement of ATP and ubiquitinating enzymes is conceptually new.

The catalytic efficiency of an enzymatic activity is determined by the Michaelis-Menten constant, K_m and the rate of catalytic turnover, k_{cat}. For Yuh1 WT, the k_{cat} of both hydrolysis and transamidation are similar, but the K_m of transamidation is an order of magnitude higher than hydrolysis. Hence, it is an efficient hydrolase than a transamidase. A close inspection of the kinetic parameters of our engineered Yuh1 mutant revealed that the k_{cat} of both hydrolysis and transamidation are heavily affected by the mutated residues. Especially for hydrolysis, the k_{cat} is two order of magnitude lower compare to the Yuh1 WT. As a result, the catalytic efficiency of transamidation is higher than hydrolysis and it manifests as a better transamidase, with the ratio of the catalytic efficiency of transamidation to hydrolysis 28-folds higher than the Yuh1 WT. Together, our engineering effort, which was a combination of targeted mutagenesis, random mutagenesis and yeast surface display-based directed evolution, has led to the emergence of a generalist with a catalytic promiscuity, which could be a starting point for the evolution of improved new function. We acknowledge that hydrolysis product is favored thermodynamically, and one of the proposed strategies for future engineering effort is to further exclude water at the catalytic cleft.

Yuh1 does not display high specificity at P1' site. In principle, it can accept any nucleophilic amine for aminolysis. Here, we have shown that the improved Yuh1 mutant can auto-ubiquitinate effectively at a Lys at the Asn164 position, which is located at the periphery of the active site. This implies that if the Lys of a substrate can be recruited to the proximity to the active site of Yuh1 for the intermolecular ubiquitination to take place.[42] Hence, our Yuh1 mutant has the potential to be developed as a modular ubiquitin transamidase. It will be interesting to fuse a substrate binding module, for example, a nanobody which recognizes our protein of interest, to the Yuh1 mutant and to investigate whether it can be ubiquitinated by the Yuh1 mutant based on the induced proximity effect.

Our study also demonstrated that Yuh1 mutant is capable to auto-ubiquitinate with defined ubiquitin chain, regardless of the linkage and length with similar efficiencies to generate anchored ubiquitin chains. We further established that the utility of these anchored ubiquitin chains as the model substrates to investigate their interactions with the ubiquitin interacting proteins, such as deubiquitinases. We have discovered that the deubiquitinases which belong to the non-linkage specific USP family, for example, USP11, USP15 and USP21 cleave the anchored ubiquitin chain from the distal end of the ubiquitin chains but they could not remove the ubiquitin conjugated to Yuh1. Surprisingly, the linkage specific OTU family, such as OTUB1 and OTUB2 are unable to process the anchored K48 linked ubiquitin chains. This observation brings into the question that the specificities of the OTU enzymes are not only solely determined by the ubiquitin linkages. Our result also supports the study that elucidates the

anchored K48 linked ubiquitin chains are resistant to the deubiquitinases activities.[43] Here, we highlighted that the nature of the ubiquitinated substrates could play an important role in regulating the activities of these deubiquitinases. Further characterization of the deubiquitinases should include the ubiquitinated substrates to fully understand the biological relevance of the deubiquitinases.

Materials and Methods

Cell culture

Saccharomyces cerevisiae strain EBY100 (MATa AGA1::GAL1-AGA1::URA3 ura352 trp1 leu2delta200 his3delta200 pep4::HIS3 prbd1.6R can1 GAL) and yeast surface display vector, pCT used in yeast display were kindly provided by Dr. Sarah Moore (Smith College).

Cloning

The *Yuh1* cDNA (IDT technologies) was cloned into pET28b (Novagen) at NdeI and XhoI (New England Biolabs) restriction site. The alanine mutants were prepared through Phusion site directed mutagenesis (New England Biolabs). For yeast display, the *Yuh1* was PCR amplified (Lucigen) using pCT-yuh1_f(2) and pCT-yuh1_r(2) (**Appendix Table S2-1**) and cloned into pCT vector at NheI and BamHI (New England Biolabs) restriction site. All primers were purchased from IDT DNA. All clonings were validated through Sanger sequencing.

Protein expression and purification

Ubiquitin (Ub) was expressed in *E. coli* Rosetta (DE3) pLysS cells using 2xYT media. To obtain 15N-labeled Ub, we supplemented M9 minimal media with ^{15}N ammonium chloride (Cambridge Isotope Laboratories). Labeled and unlabeled Ub was purified as described previously.[1] Briefly, cells were lysed in 50 mM Tris, pH 7.4, 5 mM EDTA, 1 mM EGTA, 1 mM PMSF, 1 mM DTT, 0.02% (v/v) IGEPAL and non-ubiquitin proteins were precipitated using 70% perchloric acid (1 ml for 6 L of cells). The clarified supernatant was dialyzed in 50 mM NH_4OAc, pH 4.4, 1 mM EDTA overnight before purifying through cation exchange chromatography on an ÄKTA purifier (GE Healthcare). The equilibrate buffer was 50 mM NH_4OAc, pH 4.4, 1 mM EDTA and the elute buffer was 50 mM NH_4OAc, pH 4.4, 1 mM EDTA, 1 M NaCl. For ^{15}N-labeled Ub, MALDI-TOF analysis showed that the percent of ^{15}N incorporation was 95%.

To prepare Ub-allylamine, UbD77 (final concentration was 750 μM) was added to a solution containing 50 mM HEPES pH 8, 1 mM EDTA, 100 mM allylamine (Sigma-Aldrich), and 1 μM Yuh1m. The reaction took place at room temperature for 100 min and was quenched by TFA. The reaction was buffer exchanged into 50 mM NH_4OAc, 1 mM EDTA, pH 4.4 and purified through cation exchange chromatography with the same buffer as mentioned above. Fractions containing Ub-allylamine were identified by MALDI-TOF MS.

To prepare Ub-fluorescein- or biotinylated D77, Ub K63C D77 (final concentration was 200 μM) was incubated at room temperature in a buffered solution containing 50 mM HEPES pH 7.5, 5 mM $MgCl_2$, 25 mM KCl, and 1 mM

TCEP. To this mixture was added 1 mM 5-iodoacetamidofluorescein (Sigma-Aldrich) (or 1 mM EZ-link™ iodoacetyl-PEG_2-biotin (ThermoFisher Scientific)) in 10% DMSO and the resulting mixture was incubated at 37 °C overnight. The reaction was monitored by MALDI-TOF MS and quenched with 10 mM DTT once conversion was complete. The reaction was buffer exchanged into MilliQ water before further purifying using a Zebra™ spin column, 7000 MW (Thermo Scientific).

The 6xHis-tagged Yuh1 variants and OTUB1-UBE2D2, 6XHis MBP-tagged USP11, USP15, OTUB1 and OTUB2, 6xHis SUMO-tagged USP21 were expressed in *E. coli* BL21 (DE3) or Rosetta (DE3). The cells were lysed in 50 mM sodium phosphate (or Tris), pH 8, 300 mM NaCl, 10 mM imidazole and the proteins were purified using cobalt (Agarose Bead Technologies) or nickel (Clontech) affinity chromatography. The equilibrate buffer was 50 mM sodium phosphate (or Tris), pH 8, 300 mM NaCl, 10 mM imidazole and the elution buffer was 50 mM sodium phosphate (or Tris), pH 8, 300 mM NaCl, 300 mM imidazole. For 6xHis MBP-tagged protein, TEV cleavage reaction were performed to generate tag-less protein; for 6xHis SUMO-tagged protein, ULP1 cleavage reaction was performed. These proteins were further purified through anion exchange chromatography (or size exclusion chromatography for USP21) on an ÄKTA purifier (GE Healthcare). For the anionic exchange chromatography, the equilibrate buffer is 50 mM Tris, pH 7.5, 1 mM DTT and the elution buffer is 50 mM Tris, pH 7.5, 1 mM DTT and 1 M NaCl. For the size exclusion

chromatography, the running buffer is 50 mM Tris, pH 7.5, 300 mM NaCl and 1 mM DTT.

Transamidation activity assay

Transamidation reactions were carried out in a buffered solution containing 50 mM HEPES pH 8.0, 1 mM EDTA, 100 mM allylamine (Sigma Aldrich), 750 μM Ub D77, and 1 μM enzyme. The reactions were quenched with a solution containing 25 mM iodoacetamide (IAA) and 10% (v/v) acetic acid. Reactions were monitored by MALDI-TOF MS as described below.

MALDI-TOF MS sample preparation and analysis

To 1 μl of quenched sample was added 1 μl of uniformly ^{15}N-labeled Ub D77 (5 μM). The resulting mixture was then added to 2 μl saturated sinapic acid in 40% acetonitrile, 0.1% TFA. 2 μl of this mixture was then spotted on a 96-ground steel MALDI target plate. MS spectra were acquired on a Bruker MALDI-TOF microflex. For each spectrum, we collected a total of 1000 shots at 45% laser power. Data processing was performed using FlexAnalysis software (Bruker). Calibration curves were obtained for wild type Ub and Ub allylamine to quantify the amount of hydrolysis and transamidation products.

Measuring kinetics of Yuh1 mutants using HPLC

Kinetic assays were carried out at room temperature in buffered solution containing 50 mM HEPES pH 8.0, 1 mM EDTA, 100 mM allylamine, and either

7.5 nM wild type Yuh1, or 500 nM of the Yuh1 N88A and Yuh1m variants. Initial rates were measured by varying the concentration of fluorescein-labeled Ub D77. Reactions were quenched at three different time points (0, 1, and 2 min). The fluorescein-labeled components (Ub D77, wild type Ub, and Ub allylamine) were then resolved by HPLC using a cationic exchange column (TOSOH Bioscience) and a buffer containing 25 mM NH_4OAc pH 4.4. A linear gradient of 0.15 – 0.35 M NaCl over 12 minutes was used to elute the Ub variants, which were detected by absorbance at 494 nm. Quantification was based on a calibration curve of fluorescein-labeled Ub D77. The kinetic data were analyzed using OriginPro and fit to the Michaelis-Menten equation.

Ubiquitin chain synthesis

To prepare K48 Ub_2 with UbD77 at the proximal site, 1 mM of UbD77 and 1 mM Ub (or UbK6R/K48R for initial autoubiquitination assay with Yuh1qm) were reacted with 0.2 μM E1 and 2 μM Cdc34 in a buffered solution containing 40 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM ATP, and 0.2 mM DTT at pH 7.5. The reaction was incubated at 37 ^oC overnight. For K63 Ub_2, 0.75 mM of UbD77 and 0.75 mM of Ub (or UbK11R/K63R for initial autoubiquitination assay with Yuh1qm) were reacted with 1 μM of E1, 8 μM of Ubc13-Mms2 in a buffered solution containing 40 mM Tris, 10 mM $MgCl_2$, 10 mM ATP, and 0.6 mM DTT at pH 7.5. Again, the reaction was incubated at 37^oC overnight.[2] For K6/48 Ub_3, 1 mM UbD77 and 2 mM UbK6R/K48R were reacted with 0.2 μM E1, 2 μM UbcH7, 0.45 μM NleL in a buffered solution containing 40 mM Tris, 10 mM $MgCl_2$, 10 mM

ATP, and 0.6 mM DTT at pH 7.5 for 37°C overnight. The reactions were quenched using 10 mM DTT in 50 mM NH_4OAc pH 4.4 solution for at least 30 min. Dimers were then purified using size exclusion chromatography using a HiLoad Superdex 75 pg or cationic exchange chromatography on an ÄKTA purifier. The running buffer for size exclusion chromatography was 50 mM Tris, pH 7.5, 1 mM DTT, 300 mM NaCl. For cationic exchange chromatography, the equilibrate buffer was 50 mM NH_4OAc, 100 mM NaCl while the elution buffer was 50 mM NH_4OAc and 1 M NaCl. To prepare K48 Ub_3 with proximal UbD77, two sequential Cdc34 reaction was performed. In the first reaction, 1 mM UbWT was used to generate K48 ubiquitin chains; for the second reaction, 0.5 mM of K48 Ub_2 purified from the previous reaction and 0.5 mM of UbD77 were used.

Autoubiquitination assay

Autoubiquitination reactions were performed at room temperature in a buffered solution containing 50 mM HEPES pH 8.0 and 1 mM EDTA. For mono-ubiquitination reaction, 15 μM of the Yuh1 variant (wild type, T163K, Yuh1m, or Yuh1qm) was used along with 200 μM of Ub D77; for di-ubiquitination reaction, 5 μM of Yuh1qm was reacted with 50 μM Ub chains with D77 at the proximal end. Reactions were quenched by boiling in SDS-loading dye at different time points. Western blot analysis was performed using the mouse anti-Ubiquitin (P4D1) antibody (Enzo Life Sciences) at a 1:1000 dilution. We used the IRDye® 800CW goat anti-mouse (LI-COR Biosciences) at a 1:15000 dilution as the secondary antibody.

Mass spectrometry analysis of autoubiquitination reactions

Autoubiquitination reactions were separated by SDS-PAGE using NuPAGE™ 12% Bis-Tris protein gels (ThermoFisher Scientific). The bands corresponding to mono-ubiquitinated Yuh1qm were excised and prepared for in-gel digestion with GluC at a 1:10 g/g ratio and trypsin at a 1:20 g/g ratio. K48 di-ubiquitinated Yuh1qm and K63 tri-ubiquitinated Yuh1qm5.6 was purified by a size exclusion chromatography (50 mM Tris, pH 7.5, 300 mM NaCl, 1 mM DTT) or anionic exchange chromatography (50 mM Tris, pH 7.5, 1 mM DTT as equilibration buffer and 50 mM Tris, pH 7.5, 500 mM NaCl, 1 mM DTT as elution buffer) and lyophilized. The dry sample (0.35 µg/ul) was denatured for 1 h at room temperature in a solution containing 25 mM ammonium bicarbonate and 8 M urea. Digestion of di/tri-ubiquitinated Yuh1 variants were performed under the same reaction conditions as described above.

Proteomic analysis was performed at the UMASS-Amherst Mass Spectrometry Center. Briefly, nanoLC-MS/MS was conducted in an Orbitrap Fusion™ Tribid™ mass spectrometer (ThermoFisher Scientific). Digested peptides were separated in a nanoLC column (Thermo Acclaim PepMap RSLC column, 75 µm X 15 cm) with an Easy-nLC 1000 chromatography system (ThermoFisher Scientific) using a linear gradient 0-50% buffer B (acetonitrile, 0.1% formic acid) over 50 min. The eluted peptides were analyzed with resolution of 60000 and scan range of 350-2000 m/z. Tandem mass spectrometry was performed using collision induced dissociation (CID) with a 35% collision energy. The MS data were analyzed using

Proteome Discoverer™ (ThermoFisher Scientific, Version 2.2) and SEQUEST HT search engine.

Yeast growth and induction

Yeast were transformed using the Frozen EZ Yeast Transformation II kit (Zymo Research) and plated on tryptophan-deficient dropout media (SDCAA plates; 20 g/L dextrose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.56 g/L $NaH_2PO_4.H_2O$, 15 g/L agar, 182 g/L sorbitol) for 2-4 days at 30 ^oC to identify successful transformants. Transformed EBY100 cells were grown in 5 ml of SDCAA (20 g/L dextrose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.56 g/L $NaH_2PO_4.H_2O$) media overnight at 30 ^oC. We then pelleted at least 5 X 10^7 cells and resuspended them in 5 ml of tryptophan deficient dropout media containing galactose (SGCAA media; 20 g/L galactose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.56 g/L $NaH_2PO_4.H_2O$) for induction to occur overnight at 20 ^oC.

Validation of the yeast display strategy

We harvested 5 X 10^6 cells and washed them with PBSF buffer (0.1% (v/v) BSA in PBS buffer, pH 7.4). Cells were then labeled with chicken anti-HA (1:250 dilution in PBSF buffer) (Gallus Immunotech) for 30 min at room temperature. Labeled cells were washed with PBSF buffer, aspirated, and incubated with 50 µM of biotinylated Ub D77 in 50 µl PBSF buffer for 10 min at room temperature. Reactions were quenched with 5 µM of the pan-DUB inhibitor PR-619 (Selleck

Chemicals). Since PR-619 is a reversible inhibitor, we also added it to the PBSF wash buffer. Ubiquitinated cells were simultaneously labeled with a goat, anti-chicken antibody conjugated to Alexa 647 (1:100 dilution in PBSF buffer) (Invitrogen™) and streptavidin conjugated to Alexa 488 (1:100 dilution in PBSF buffer) (Invitrogen™) for 15 min on ice in a shielded box. Cells were washed one last time with PBSF buffer and resuspended in the same buffer prior to performing flow cytometry on a BD LSR Fortessa X20. We counted cells based on both high expression levels of Yuh1 and high autoubiquitination efficiency.

Construction of the Yuh1qm library

The Yuh1qm gene was mutagenized using Mutazyme II, a commercially available error prone polymerase (Agilent Technologies). The primers were the same as those used to amplify the DNA library pool: pCT-yuh1_f2 and pCT-yuh1_r2 (**Appendix Table S2-1**). To prepare electrocompetent yeast cells, a 25 ml culture of EBY100 cells were grown to OD 1.5 and treated with 25 mM DTT for 15 min. Cells were then harvested, washed, and resuspended in 150 µl E buffer (10 mM tris, pH 7.5, 270 mM sucrose, 2 mM $MgCl_2$). Cells were transformed with 1 µg of the NheI and BamHI digested plasmid and 4 µg of amplified DNA library pool through electroporation. Transformed cells were serial diluted and plated in -Trp dropout media for 2-4 days to estimate the size of the library.

Selection of Yuh1qm library displayed on the yeast surface

Five rounds of selection were carried out using fluorescent activated cell sorting (FACS) on a BD FACS ARIA II with 488 nm (530/30 band pass collection filter) and 640 nm (670/30 band pass filter) excitation lasers. For the first round of selection, 2 X 10^8 cells displaying the Yuh1qm library were harvested and washed with PBSF buffer. Aspirated cells were then incubated with biotinylated Ub D77 using the procedure described above. In the second round, 50 μM of biotinylated Ub D77 was incubated with 2 X 10^7 cells. In subsequent rounds, 5 X 10^6 cells were harvested and reacted with 5 μM, 0.5 μM, and 0.1 μM of biotinylated Ub D77. The reaction time was 10 min for all selection rounds except round 5, which was 5 min. At the end of rounds three and five, plasmids were isolated from the enriched clones using a Zymoprep yeast plasmid miniprep kit (Zymo Research) and electroporated into DH10β cells. Individual colonies were grown, miniprepped and sequenced using pCT_yuh1_r2 to identify beneficial mutations for autoubiquitination.

Validation of selected clones through autoubiquitination assays

Selected clones were amplified using the primers Yuh1_for_NdeI and Yuh1_rev_XhoI, cloned into pET28b, expressed, and purified using the procedure described above. For autoubiquitination, 1 μM of Yuh1qm variants was incubated with 2 μM Ub D77 in 50 mM HEPES, pH 8.0, 1 mM EDTA. Reaction progress was monitored by gel densitometry; gels were stained using

SYPRO™ Ruby (Invitrogen™) and scanned on a Typhoon FLA 9500 (GE Healthcare). Data analysis was performed on Image Studio Lite (LI-COR Biosciences). Data corresponding to the hydrolysis of the autoubiquitinated species were fit to an exponential decay curve on OriginPro.

DUB assays with anchored and unanchored Ub chains

For maximum conversion of free ubiquitin chains into ubiquitinated Yuh1qm5.6, 5 μM of clone 5.6 was incubated with 2.5 μM ubiquitin chains (with Ub D77 as the proximal unit) for 30 s. Reactions were quenched with 1 mM NEM and buffered exchanged into MilliQ water. The resulting solutions were used directly in DUB assays. For DUB assays, 0.5 μM of the USPs or 1 μM of OTUs were incubated with the free and anchored ubiquitin chains. Reaction progress was monitored by gel densitometry. For free ubiquitin chains, gels were stained with SYPRO™ Ruby or transferred and blotted with the P4D1 anti-ubiquitin antibody (1:1000 dilution in 3% BSA, 0.1 % TWEEN; Enzo Life Sciences). For anchored ubiquitin chains, gels were transferred and blotted with anti-6X His tag® antibody (1:1000 dilutions in 3% BSA, 0.1% TWEEN; Abcam). All western blots were imaged on Odyssey CLx imaging system (LI-COR biosciences). Data analysis was performed on Image Studio Lite.

References

1. Chang, L. H. & Strieter, E. R. Reprogramming a Deubiquitinase into a Transamidase. *ACS Chem. Biol.* **13**, 2808–2818 (2018).
2. Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).

3. Avram Hershko, A. C. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425 (1998).

4. Komander, D. & Rape, M. The Ubiquitin Code. *Annu. Rev. Biochem.* **81**, 203–229 (2012).

5. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).

6. Komander, D., Clague, M. J. & Urbé, S. Breaking the chains: Structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**, 550–563 (2009).

7. Mevissen, T. E. T. & Komander, D. Mechanisms of Deubiquitinase Specificity and Regulation. *Annu. Rev. Biochem.* **86**, 159–192 (2017).

8. Cao, J. & Yan, Q. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Front. Oncol.* **2**, 1–9 (2012).

9. Pickart, C. M. Targeting of substrates to the 26S proteasome. *FASEB J.* **11**, 1055–1066 (1997).

10. Uckelmann, M. & Sixma, T. K. Histone ubiquitination in the DNA damage response. *DNA Repair (Amst).* **56**, 92–101 (2017).

11. Xia, Z. P. *et al.* Direct activation of protein kinases by unanchored polyubiquitin chains. *Nature* **461**, 114–119 (2009).

12. Clague, M. J., Heride, C. & Urbé, S. The demographics of the ubiquitin system. *Trends Cell Biol.* **25**, 417–426 (2015).

13. Dong, K. C. *et al.* Preparation of distinct ubiquitin chain reagents of high purity and yield. *Structure* **19**, 1053–1063 (2011).

14. Martinez-Fonts, K. & Matouschek, A. A Rapid and Versatile Method for Generating Proteins with Defined Ubiquitin Chains. *Biochemistry* **55**, 1898–1908 (2016).

15. Martinez-Fonts, K. *et al.* The proteasome 19S cap and its ubiquitin receptors provide a versatile recognition platform for substrates. *Nat. Commun.* **11**, (2020).

16. Saeki, Y., Isono, E. & Toh-E, A. Preparation of ubiquitinated substrates by the PY motif-insertion method for monitoring 26S proteasome activity. *Methods Enzymol.* **399**, 215–227 (2005).

17. Chatterjee, C., McGinty, R. K., Pellois, J. P. & Muir, T. W. Auxiliary-mediated site-specific peptide ubiquitylation. *Angew. Chemie - Int. Ed.* **46**, 2814–2818 (2007).

18. McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G. & Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453**, 812–816 (2008).

19. Johnston, S. C., Riddle, S. M., Cohen, R. E. & Hill, C. P. Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J.* **18**, 3877–3887 (1999).

20. Valkevich, E. M. *et al.* Forging isopeptide bonds using thiol-ene chemistry: Site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases. *J. Am. Chem. Soc.* **134**, 6916–6919 (2012).

21. Trang, V. H., Rodgers, M. L., Boyle, K. J., Hoskins, A. A. & Strieter, E. R. Chemoenzymatic synthesis of bifunctional polyubiquitin substrates for monitoring ubiquitin chain remodeling. *ChemBioChem* **15**, 1563–1568 (2014).

22. Yu, H. A. *et al.* Characterization of ubiquitin C-terminal hydrolase 1 (YUH1) from Saccharomyces cerevisiae expressed in recombinant Escherichia coli. *Protein Expr. Purif.* **56**, 20–26 (2007).

23. Ritorto, M. S. *et al.* Screening of DUB activity and specificity by MALDI-TOF mass spectrometry. *Nat. Commun.* **5**, (2014).

24. Khersonsky, O. & Tawfik, D. S. Enzyme Promiscuity : A Mechanistic and Evolutionary Perspective. (2010) doi:10.1146/annurev-biochem-030409-143718.

25. Copley, S. D. Shining a light on enzyme promiscuity. *Curr. Opin. Struct. Biol.* **47**, 167–175 (2017).

26. De Bie, P. & Ciechanover, A. Ubiquitination of E3 ligases: Self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms. *Cell Death Differ.* **18**, 1393–1402 (2011).

27. Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci.* **108**, 11399–11404 (2011).

28. Boder, E. T. & Wittrup, K. D. Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol.* **328**, 430–444 (2000).

29. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. **1**, 755–769 (2007).

30. Chen, T. F., De Picciotto, S., Hackel, B. J. & Wittrup, K. D. *Engineering fibronectin-based binding proteins by yeast surface display*. *Methods in Enzymology* vol. 523 (Elsevier Inc., 2013).

31. Jan Deniau, A. H., Subileau, M. & Dubreucq, E. Characterization and Reshaping of a Large and Hydrophobic Nucleophile Pocket in Lipases/Acyltransferases. *ChemBioChem* **19**, 1839–1844 (2018)

32. Harper, S. *et al.* Structure and catalytic regulatory function of ubiquitin specific protease 11 N-terminal and ubiquitin-like domains. *Biochemistry* **53**, 2966–2978 (2014).

33. Crowe, S. O. *et al.* Subunit-Specific Labeling of Ubiquitin Chains by Using Sortase: Insights into the Selectivity of Deubiquitinases. *ChemBioChem* 1525–1531 (2016) doi:10.1002/cbic.201600276.

34. Ye, Y. *et al.* Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21. *EMBO Rep.* **12**, 350–357 (2011).

35. Edelmann, M. J. *et al.* Structural basis and specificity of human otubain 1-mediated deubiquitination. *Biochem. J.* **418**, 379–390 (2009).

36. Altun, M. *et al.* The human otubain2-ubiquitin structure provides insights into the cleavage specificity of poly-ubiquitin-linkages. *PLoS One* **10**, 1–15 (2015).

37. Mevissen, T. E. T. *et al.* OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell* **154**, 169–184 (2013).

38. Wiener, R. *et al.* E2 ubiquitin-conjugating enzymes regulate the deubiquitinating activity of OTUB1. *Nat. Struct. Mol. Biol.* **20**, 1033–1039 (2013).

39. Hospenthal, M. K., Freund, S. M. V. & Komander, D. Assembly, analysis and architecture of atypical ubiquitin chains. *Nat. Struct. Mol. Biol.* **20**, 555–565 (2013).

40. Wiener, R., Zhang, X., Wang, T. & Wolberger, C. The mechanism of OTUB1-mediated inhibition of ubiquitination. *Nature* **483**, 618–622 (2012).

41. Schulman, B. A. & Wade Harper, J. Ubiquitin-like protein activation by E1 enzymes: The apex for downstream signalling pathways. *Nat. Rev. Mol. Cell Biol.* **10**, 319–331 (2009).

42. Portnoff, A. D., Stephens, E. A., Varner, J. D. & DeLisa, M. P. Ubiquibodies, synthetic E3 ubiquitin ligases endowed with unnatural substrate specificity for targeted protein silencing. *J. Biol. Chem.* **289**, 7844–7855 (2014).

43. Schaefer, J. B. & Morgan, D. O. Protein-linked ubiquitin chain structure restricts activity of deubiquitinating enzymes. *J. Biol. Chem.* **286**, 45186–45196 (2011).

CHAPTER 3

ENGINEER A NANOBODY SPECIFIC TOWARDS PROTEASOME-ASSOCIATED UCH37

The work described in this Chapter is being prepared for manuscript submission.

Abstract

In eukaryotes cell, proteasome is the main protein degradation machinery through ubiquitin proteasome system, and the proteasome-associated deubiquitinases (DUBs) play an important role in regulating the activities of proteasome by hydrolyzing the ubiquitin conjugated to the proteasome substrate. Here, we report the engineering of nanobodies targeting UCH37, which is a proteasome-associated deubiquitinase. Through a yeast display-based directed evolution using ubiquitin conjugated UCH37 complex and nonpurified proteasome associated UCH37 as antigen, we have discovered nanobodies with sub-nanomolar affinities towards UCH37. These nanobodies also interact tightly with recombinant UCH37 in complex with RPN13, which is a proteasome associated subunit that recruits UCH37. We further demonstrated that one of our engineered nanobody immunoprecipitates with the endogenous UCH37 containing proteasome *in cellulo*. We envision that these nanobodies could be developed as intracellular tools to probe UCH37 spatially and temporally in a cellular context.

Introduction

In living system, the protein homeostasis is regulated through a delicate balance between protein synthesis and protein degradation. For eukaryotes, the ubiquitin-proteasome system (UPS) is the main protein degradation machinery, which is responsible of 80% of the protein degradation.[1] 26S Proteasome is a 2.6 MDa protein complexes, which can be categorized into 19S regulatory particle (19S RP) and 20S core particle (20S CP). In the UPS, the proteasome substrate is ubiquitinated through an ATP dependent enzymatic cascade, involving E1 activating enzymes, E2 conjugating enzymes, and E3 ligases. The proteasomal ubiquitin receptors will recognize the ubiquitinated substrate, which lead to the recruitment of proteasome to the substrate. The ubiquitin on the substrate will be cleaved and removed by the proteasome-associated deubiquitinases (DUBs) locating at the 19S RP, which are UCH37, USP14 and RPN11. The substrate will then be unfolded by the ATPases, also situated at 19S RP. Next, the unfolded substrate will be degraded by the proteolytic chamber in the 20S CP.[2,3,4,5]

Every event in the ubiquitin proteasome system are regulated to ensure the proper function of a proteasome. Here, we are interested in studying the role of a proteasome associated deubiquitinase, UCH37. As suggested from its name, UCH37 is a 37 kDa cysteine protease which belongs to the ubiquitin C-terminal hydrolase (UCH) family of deubiquitinases. Studies have elucidated that UCH37 is recruited to proteasome through a 19S RP proteasome subunit, RPN13, which association activates the activity of UCH37 (**Figure 3-1A**).[6,7,8] Structural studies of UCH37 and UCH37 in complex with RPN13 have shown

that UCH37 consists of a N terminal catalytic domain, which has a similar folding motif with the other UCH members; and an extended ULD domain at the C terminal (**Figure 3-1B**).[9,10] Further studies has indicated that the ULD domain is highly flexible, and is bound to the DEUBAD domain of RPN13. This interaction changes the orientation of ULD domain relative to the catalytic domain of UCH37, effectively "lock" the UCH37 in an active conformation.[11,12] Proteasome-associated UCH37 has been suggested to be an anticancer drug target and the understanding of the regulation on UCH37 is crucial in the therapeutic development.[13,14]

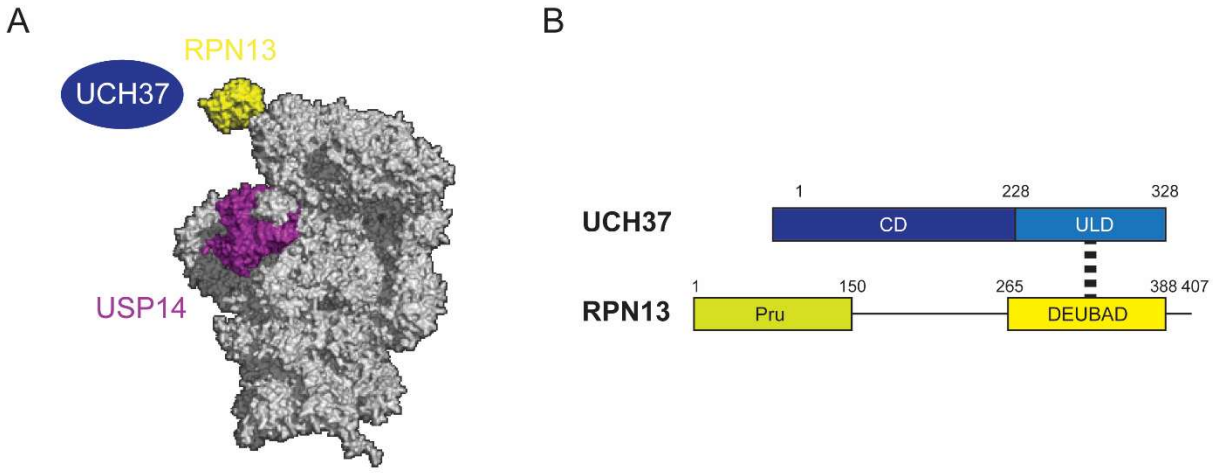

Figure 3-1: Proteasome-associated UCH37

(**A**) 3D reconstruction of yeast 26S proteasome-Ubp6 (homologous to human USP14) complex (PDB ID: 5A5B).[15] Cryo-EM structure is used as a representative scheme to depict the orientations of USP14 and UCH37 on proteasome. (**B**) UCH37 interacts with RPN13 at the C terminal ULD domain of UCH37 and the C terminal DEUBAD domain of RPN13.

The paradigm of study the function of a protein is through genetic experiment, which result in irreversible modification of the target protein. On

other hand, intrabodies or intracellular antibodies have emerged as an alternative for probing the biological functions of protein *in cellulo* and *in vivo*.[16,17] Here, we proposed to utilize nanobody as the next generation tool to investigate the proteasome associated UCH37 in a cellular context. Nanobody, or V_HH antibody, is the single monomeric variable antibody domain of the heavy chain only antibody, which is found in the *Camelidae* family (**Figure 3-2A**). Nanobody has high stability, and it recognizes its epitopes using its three complementary determining regions (CDRs), which is similar to the antigen-interacting surface of an antibody (**Figure 3-2B**). In contrast to the conventional antibody, which is composed of two heavy and light chains linked by disulfide bond with size approximate to 150 kDa, nanobody is monomeric, with molecular weight around 15 kDa. Hence, it could be genetically expressed and folded properly inside living cells. Together, using nanobody as an intrabody would enable us to probe the protein of interest in real time without genetic manipulation on the endogenous protein.[18,19]

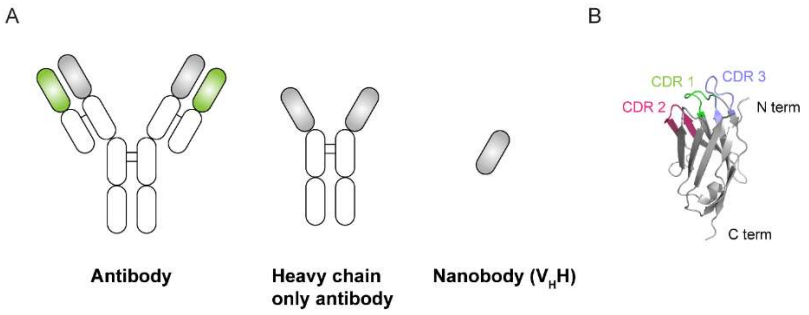

Figure 3-2: Nanobody

(**A**) Comparison of conventional antibody, heavy chain only antibody and nanobody. The variable domain of the heavy chain (gray) and light chain (green). (**B**) Crystal structure of a nanobody (PDB ID: 5VNV) with its complementary determining regions (CDRs) highlighted.

Nanobodies are generally produced through the immunization of llamas (or camel) using the antigen protein. The immunized library is then displayed on the phage surface followed by biopanning to isolate the potential nanobody binders.[19] Meanwhile, studies have shown that synthetic nanobody library display on the yeast surface has been successfully used to identify nanobody binders towards a variety of proteins. The *in vitro* screening conditions offer by synthetic nanobody libraries have enabled the discovery of conformational selective nanobodies and nanobodies targeting challenging proteins, such as membrane proteins.[20,21]

Here, we decided to engineer a nanobody specific towards proteasome-associated UCH37. By using a synthetic nanobody library displayed on yeast surface and ubiquitin conjugated UCH37·RPN13$_{DEU}$ at its catalytic Cys as antigen, we have isolated several nanobody binders, probably targeting the exosite of UCH37. We further performed affinity maturation using random mutagenesis and nonpurified proteasome associated UCH37 as antigen and have identified nanobodies with sub-nanomolar affinities towards free UCH37 and interact tightly with UCH37·RPN13. We also demonstrated that one of the isolated nanobody immunoprecipitates with endogenous UCH37 containing proteasome *in cellulo*. Our result suggested that these nanobodies could be developed as the UCH37 intrabodies to probe proteasome-associated UCH37 in a cellular context.

Result

Generation of the UCH37 nanobodies

Using a synthetic nanobody library (**Appendix A**), we designed a screening strategy to isolate the nanobodies binders for UCH37. The straightforward approach is to use purified UCH37 protein as the antigen and to select for the clones which bind to the exposed surfaces of UCH37. Nonetheless, manipulation of the full length UCH37 protein has been problematic; free UCH37 is prone to dimerize and further oligomerize in micromolar range *in vitro*.[22] Furthermore, these UCH37 oligomers adopt an auto-inhibition conformation, with one of the helixes of a subunit blocks the ubiquitin binding site of the other subunit.[10] As mentioned in the introduction section, UCH37 interacts with the DEUBAD domain of RPN13 ($RPN13_{DEU}$), which will disrupt the dimerization of UCH37, leading to the formation of UCH37·$RPN13_{DEU}$ heterodimer which activates the catalytic activities of UCH37 towards ubiquitin-7-amido-4-methylcoumarin (UbAMC).[8,11,12] Our group further showed the UCH37·$RPN13_{DEU}$ is more reactive in hydrolyzing Lys48 linked branched ubiquitin chains. As our goal is to identify nanobodies targeting proteasome bound UCH37, purified UCH37·$RPN13_{DEU}$ is the biological relevant protein to be used as the antigen for our directed evolution scheme.

Another consideration is that most deubiquitinases shared a similar and conserved S1 site to recognize ubiquitin; these enzymes interacts extensively with the C terminus of ubiquitin at the S1 binding site.[23,24] It is a well-known observation that directed evolution has the tendency to isolate binders targeting

the substrate binding site which has a well-evolved binding cleft. Hence, one of the concerns of our screening strategies is the generation of low specificity nanobodies targeting the conserved S1 ubiquitin binding site. This motivated us to prepare a UCH37·RPN13$_{DEU}$ conjugated to the inhibitor ubiquitin propargylamine (Ub-PA) at its catalytic Cys88 as the antigen for our screening (**Figure 3-3A**). We hypothesized that using the UCH37·RPN13$_{DEU}$ with the blocked S1 site will enhance our chance of identifying exosite binders with improved specificity. We recognized that the caveat of utilizing protein complex as antigen is that we might selecting for nanobodies targeting the non-desirable epitope, which are RPN13$_{DEU}$ and ubiquitin; negative selection could be performed to circumvent this issue and careful characterization need to be conducted to ensure the binding properties of the selected nanobodies.

Our initial yeast display selection scheme included two rounds of magnetic activated cell sorting (MACS) followed by two rounds of fluorescent activated cell sorting (FACS) using the biotinylated ubiquitin conjugated UCH37·RPN13$_{DEU}$. The biotin was introduced as a mean to detect our antigen in the sorting. Here, we adopted the established AviTag technology to biotinylate site specifically at the N terminus of ubiquitin.[25] In each round, a minimum of 10X of oversampling of the library size was used to avoid the loss of the unique clones. Due to its high screening capacity, which is close to 10^{10} of cells and its ability to bind to even the weak binders in micromolar range, MACS was performed at first to reduce the size of our initial library, allowing subsequent rounds to be carried out using FACS.[26] For FACS, we increased the screening stringency by lowering the

concentration of the antigen with each successive rounds. We observed the emergence of a tight binding populations at the end of FACS (**Figure 3-3B**). Sanger sequencing revealed that there was an overwhelming enrichment of a single clone, we named it as NbAI.1.

To identify additional nanobodies targeting other interaction surfaces of UCH37, we screened the sorting library against UCH37·RPN13$_{DEU}$ in complex with NbAI.1. By blocking the epitope of UCH37 recognized by NbAI.1, we reasoned this strategy will bias our screening result in isolating nanobodies either bind outside the S1 ubiquitin and NbAI.1 binding site or with improved binding kinetics that could compete of NbAI.1 (**Figure 3-3C**). Similarly, we prepared our biotinylated antigen by fusing the AviTag at the N terminus of UCH37. The screening led to the discovery of two different nanobodies, herein referred as NbAII.1 and NbAII.2 (**Figure 3-3D**). Based on phylogenetic analysis, the identified NbAI.1, NbAII.1 and NbAIII.1 are distinct (**Figure 3-4**). Interestingly, the CDR2 of NbAII.1 has one additional amino acid, which could be a result of serendipity or synthesis error during preparation of the synthetic library. We further interrogated the binding of the identified nanobodies towards UCH37·RPN13$_{DEU}$ based on the titration of nanobodies immobilized on yeast surface (**Figure B-4**); NbAII.2 was not further studied due to its unsatisfactory binding profile.[27,28]

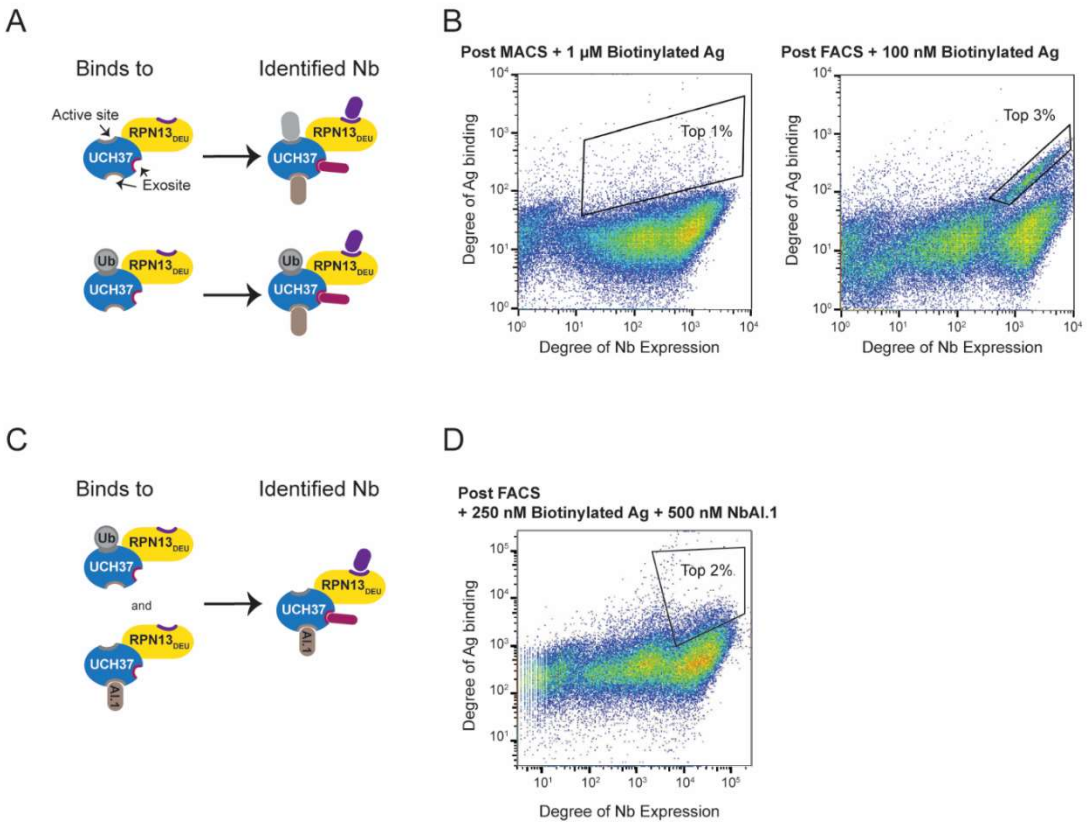

Figure 3-3: Initial screening strategy for the isolation of UCH37 nanobodies

(**A**) Scheme showing the antigen design and (**B**) FACS result in the screening against ubiquitin conjugated UCH37·RPN13$_{DEU}$. (**C**) Scheme showing the antigen design and (**D**) FACS result in the screening against UCH37·RPN13$_{DEU}$ in complex with NbAl.1. The frames and the percentages designate the fraction of yeast population isolated in each round of sorting.

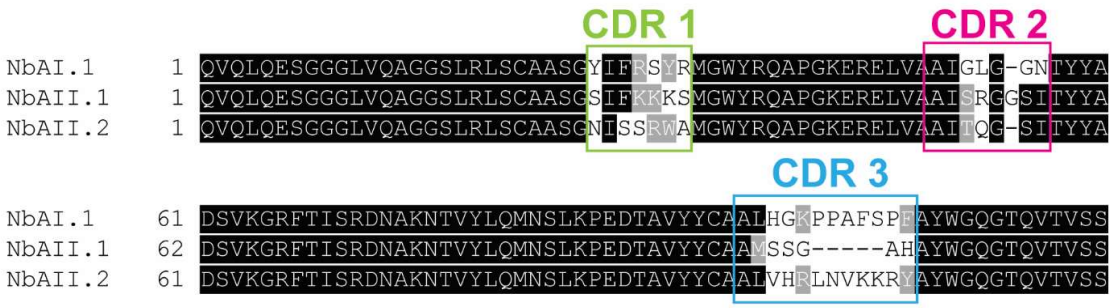

Figure 3-4: Amino acid sequences of the enriched clones

The diversified residues in the CDRs are highlighted in colored box.

Affinity maturation of the proteasome-associated UCH37 specific Nbs

As our goal is to utilize the nanobody as an intracellular probe towards proteasome associated UCH37, our ideal nanobody should recognize the proteasomal UCH37 in a cellular context. Here, we performed affinity maturation to improve the binding affinities of our identified nanobodies. As mentioned, proteasome is a multi-protein complexes consist of around 30 subunits. Studies have shown that the constitutive proteasome subunits could be fused to affinity tag to facilitate the purification of proteasome from the mammalian cell.[29,30] Especially affinity tagged RPN11 has been extensively used to purified the UCH37 containing proteasome. Meanwhile, study has shown that yeast surface Aga1p-Aga2p display platform is compatible for cell lysate-based antigen.[31] Hence, we have devised a new yeast display scheme using the cell lysate of a stable HEK293 cell line expressing RPN11-His_6-TEV-Biotin-His_6 ($HEK293^{RPN11\text{-}HTBH}$) as our antigen in the positive screening step. To prevent the enrichment of nanobodies that interact with the random cellular proteins or other proteasomal subunits in the cell lysate, our selection scheme also incorporated a negative selection scheme using the cell lysate of a CRISPR/Cas9 mediated UCH37 knockout $HEK293^{RPN11\text{-}HTBH}$ (ΔUCH37 $HEK293^{RPN11\text{-}HTBH}$) as the antigen (**Figure 3-5A**). By alternating between the positive screening and negative selection, we hypothesized this will ensure the isolation of nanobodies binding specifically to proteasome-bound UCH37 in a native environment.

In our initial screening, NbAI.1 showed the highest binding affinity in the titration experiments. Therefore, we generated a random mutagenized NbAI.1

library through error prone PCR. Upon transformation, we obtained around $2x10^7$ clones. Further inspection of our random mutagenized library through next generation sequencing yield 338904 raw reads, with 91% of the ensembles are in frame. Among the library members, 31% are unique with 17% represent the original NbAl.1 (**Table 3-1**). As a result, the effective size of our random mutagenized NbAl.1 library is estimated to be around $6x10^6$ variants. A random sampling of 500 clones suggested that there are 1-7 mutations in the NbAl.1 variants. We further validated the feasibility of our screening strategies by incubating the induced library with the WT and UCH37 KO HEK293$^{RPN11\text{-}HTBH}$ cell lysate and detected a visible difference in the degree of antigen binding (**Figure B-5**).

Table 3-1: Overview of sequencing data of random mutagenized NbAl.1 library

DESCRIPTION	NO. OF READS	PERCENTAGE (%)
RAW	338904	100.0
IN FRAME	307933	90.9
UNIQUE	105222	31.0
NBAl.1	58406	17.2

Raw data represents the total number of merged reads with low quality data removed; in frame is the number of reads containing in frame nanobody framework; unique read is the number of nanobody with unique sequence; and NbAl.1 is the number of reads with no mutation.

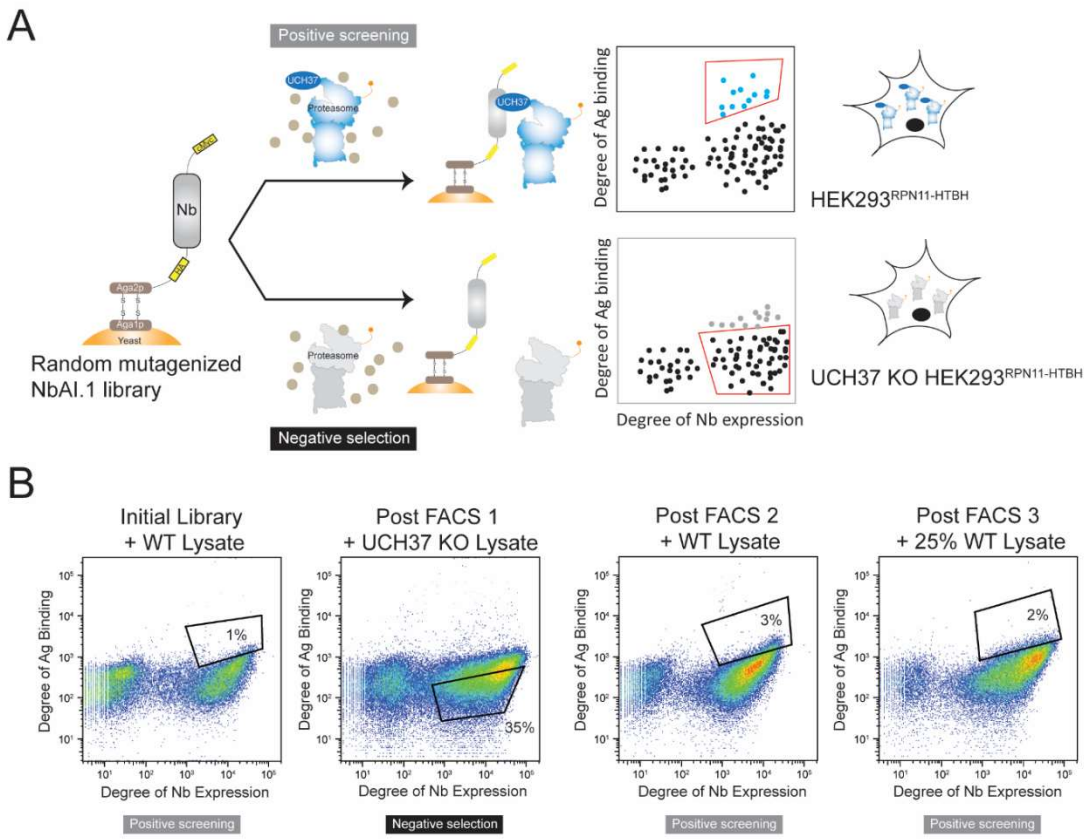

Figure 3-5: Affinity maturation of the proteasome-associated UCH37 nanobodies

(**A**) Positive screening and negative selection scheme and (**B**) FACS result using HEK293$^{RPN11\text{-}HTBH}$ cell lysate. The frames and the percentages designate the fraction of yeast population isolated in each round of sorting.

Next, we performed our enrichment strategy through FACS; we started by one round of positive screening, followed by one round of negative selection and two rounds of subsequent positive screenings with diluted cell lysate (**Figure 3-5B**). Upon sorting, the size of the isolated clones is reduced to be 10^4 clones. The gene of the pre- (the random mutagenized NbAl.1 library) and post-selection clones were pooled together and subjected to multiplex next generation sequencing to identify the enriched variants. The number of reads assigned for

the post- selection clones were in the order of 10^5 to ensure the all the isolated nanobodies were sequenced. Not surprisingly, with close to 20% of NbAI.1 presence in the random mutagenized library, the clones with the highest frequencies of the post-selection clones are NbAI.1 (**Table 3-2**). This explained why we did not discover any clones with mutations was being enriched in our screening through sanger sequencing. To identify the beneficial mutations, we calculated the enrichment ratio of the clones based on the following equation:[32]

$$Enrichment\ ratio = \frac{Normalized\ frequency\ of\ clone_{post-selection}}{Normalized\ frequency\ of\ clone_{pre-selecti}},$$

The frequency of pre-selection clone was set at minimum to one.

It is crucial to note that only around $4x10^5$ reads was assigned to the pre-selection nanobodies, which was far less than the size of the transformed library (only 2%) and that a great portion of nanobodies were not sequenced. Hence, for calculation purpose, the frequency of the pre-selection clones was set to one at minimum.

Table 3-2: Overview of the sequencing data of the post selection clones

DESCRIPTION	NO. OF READS	PERCENTAGE (%)
Raw	124883	100.0
NbAI.1	23800	19.1

Raw data represents the total number of merged reads with low quality data removed and NbAI.1 is the number of reads with no mutation.

Our result indicated that the top 20 enriched clones showed the enrichment ratio in the range of 100~400 (**Figure 3-6A**). Especially the top three clones have consensus mutations at three positions, they are Q1R (or K), H99Y

and A109P (**Figure 3-6B**). Moving down the list revealed another enriched mutation at R33P. Not surprisingly, most of the consensus mutations, R33P, H99Y and A109P are located at the CDRs, which are the regions responsible for antigen recognition. We further mapped those mutated residues on a solved crystal structure of nanobody with 85% of sequence identity (PDB ID: 6O3C); it was suggested that these residues are clustered (**Figure 3-6C**).[33]

We further displayed these enriched NbAI.1 variants on the yeast surface and interrogated their binding towards the cell lysate generated from HEK293$^{RPN11\text{-}HTBH}$. Compare to NbAI.1 and NbAI.1 variant without the enriched mutations, variants containing single mutation at Q1K, R33P, H99Y and A109P showed higher degree of binding, supporting our NGS data that these mutations are beneficial (**Figure 3-7**). We did not observe improved binding with the incubation with the UCH37 KO HEK293$^{RPN11\text{-}HTBH}$, suggesting that these nanobodies are specific towards proteasome-associated UCH37 (**Figure B-6**). Here, NbAIII.15 (Q1R/L80Q/H99Y/A109P) and NbAIII.16 (Q1R/H99Y) which combine the beneficial mutations were chosen for further studies. We would like to emphasize that L80Q was a random mutation that was originated from the cloning using the plasmid extracted from a sorted clone and was not a mutation identified in our NGS data. In addition, we observed that the recombinantly expressed nanobody containing R33P consistently precipitated during protein purification and was not further studied.

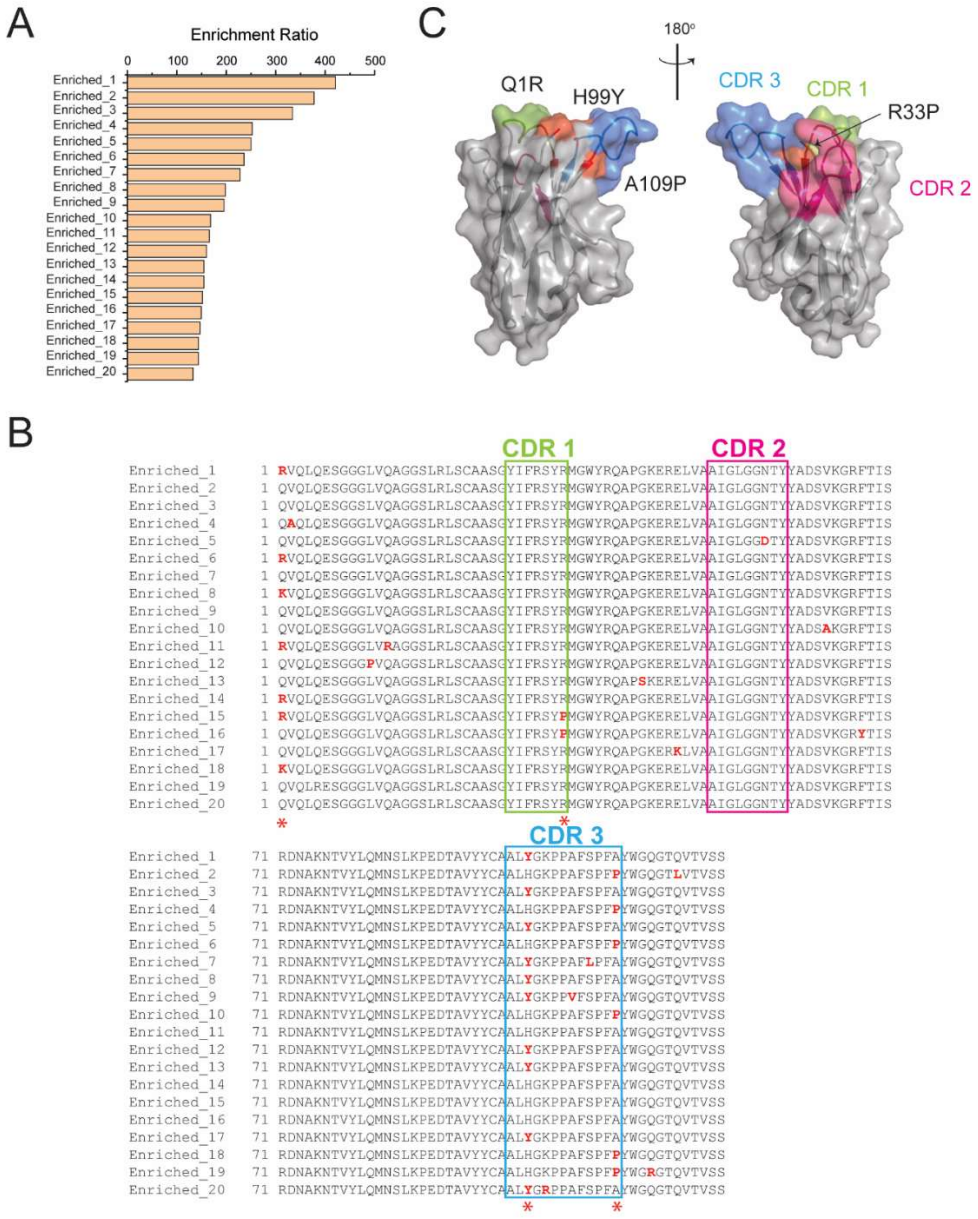

Figure 3-6: Enriched NbAl.1 variants

(A) Enrichment ratio and (**B**) the sequence alignment of the top 20 enriched clones. The CDRs are featured in the colored box. The mutations are printed in red and the consensus mutations are highlighted by asterisk (*). (**C**) Crystal structure of nanobody (PDB ID: 6O3C) with enriched mutations colored in red.

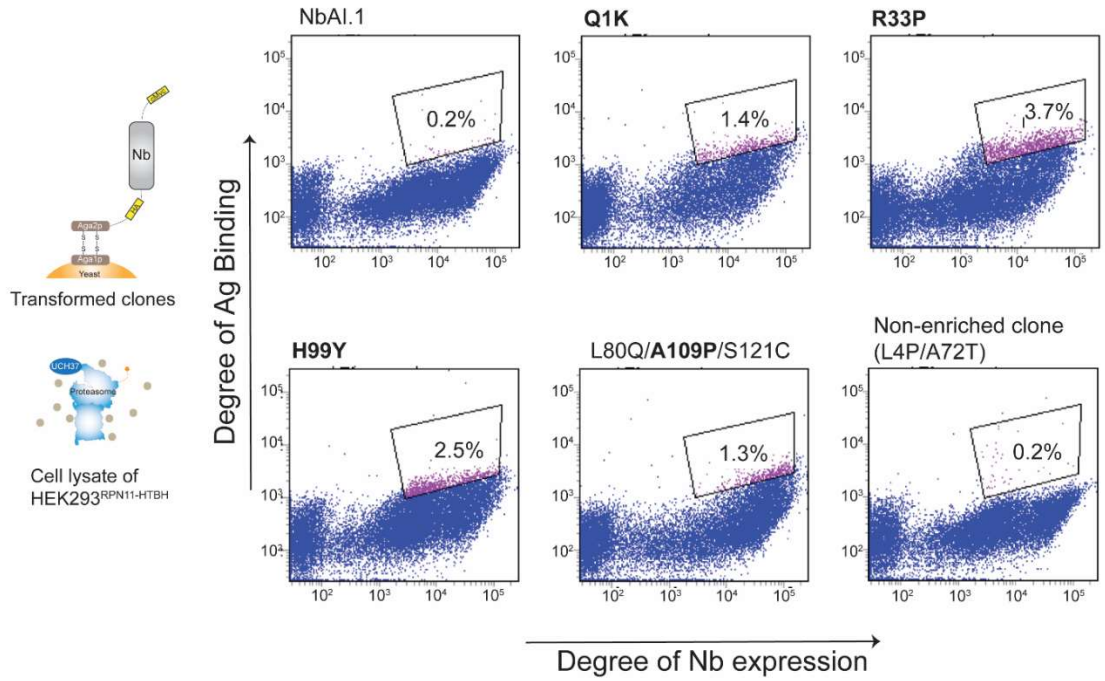

Figure 3-7: Investigation of the enriched mutations

Flow cytometry plots showing the binding of the yeast surface displayed-nanobodies containing the enriched mutations towards the cell lysate generated from the HEK293$^{RPN11\text{-}HTBH}$. The enriched mutations are highlighted in bold. The frames designate the percentages fraction of yeast population with improved degree of binding.

Characterization of Nanobodies

For the characterization of the protein-protein interactions (PPIs) of our identified nanobodies with proteasomal UCH37, we performed the biophysical characterization of these nanobodies in different stages. We started by accessing the binding of the nanobodies with UCH37 qualitatively through a pull-down assay. Here, we expressed the nanobodies as a HaloTag fusion construct and incubated these nanobodies with the recombinantly expressed UCH37 and UCH37·RPN13. Our result elucidated that all nanobodies can capture the purified

UCH37 and UCH37·RPN13, although the SDS-PAGE analysis suggested that the interaction between NbAII.1 with UCH37·RPN13 was apparently weaker (**Figure 3-8A and B**). We recognized that the concentration of the full length UCH37 used in this *in vitro* pull-down assay was at the micromolar range, when the UCH37 exists in the dimeric or oligomeric status; the result did not provide insight in how the UCH37 polymers interacts with the nanobodies.

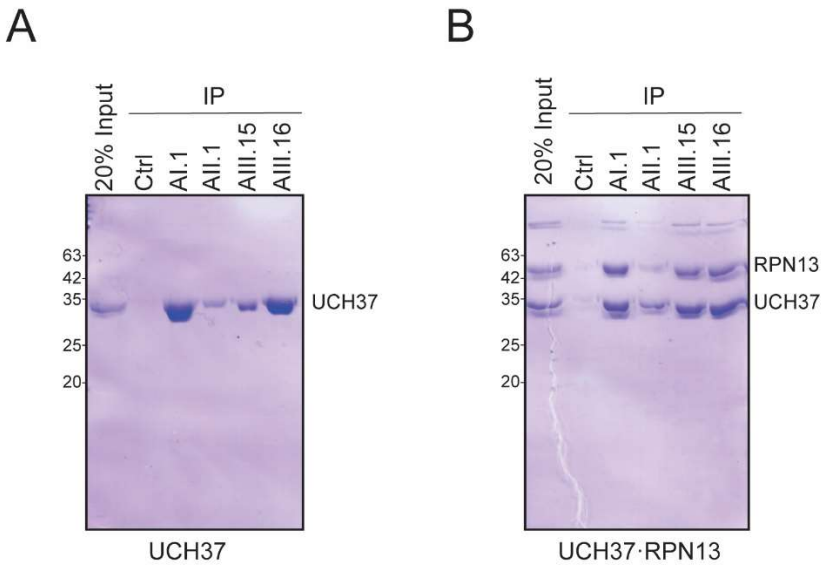

Figure 3-8: Pull-down assay with recombinant UCH37

In vitro pull-down assay of HaloTag nanobodies with recombinantly expressed (**A**) full length UCH37 and (**B**) UCH37·RPN13. Ctrl is a random nanobody which does not interact with UCH37 and its binding partner.

Next, we characterized the bindings of the nanobodies with free UCH37 quantitatively through surface plasmon resonance (SPR). Here, to ensure the binding measurement was between the nanobodies and free UCH37, UCH37 in nanomolar range was flowed over the immobilized nanobodies. Our result shown that these nanobodies are tight binders to UCH37, with K_D ranges from low to

sub-nanomolar affinities (**Figure 3-9**). NbAI.1 and its derivatives, NbAIII.15 and NbAIII.16 bind to UCH37 at 0.3±0.1 nM, 0.13±0.02 nM and 0.21±0.05 nM respectively, while the K_D of NbAII.1 is 5.6±0.8 nM (**Table 3-3**). Our result also indicated that for NbAIII.15 and NbAIII.16, the improvements in the binding are mainly derived from the slower dissociation rate.

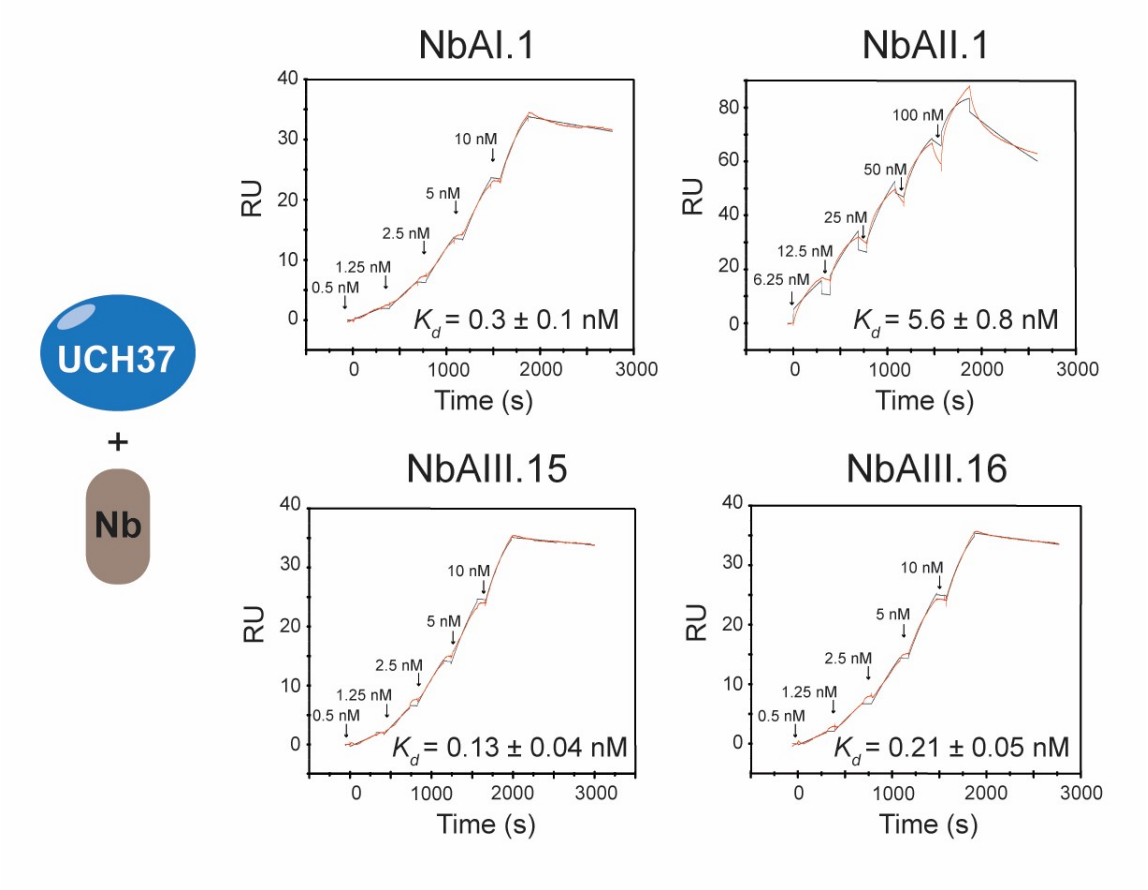

Figure 3-9: Kinetic analysis of the binding of nanobodies with free UCH37

Representative double referenced sensorgrams (red) of full length UCH37 binding to immobilized NbAI.1 (Top left), NbAII.1 (Top right), NbAIII.15 (Bottom left) and NbAIII.16 (Bottom right). Binding experiments were performed in single-cycle kinetic format at 25°C.[34] Black line represents the fitted sensorgrams.

Table 3-3: Binding parameters of the UCH37 nanobodies

Variants	k_{on} (x10^5 $M^{-1}s^{-1}$)	k_{off} (x10^{-4} s^{-1})	K_D (nM)
NbAI.1	3.9±0.6	1.3±0.4	0.3±0.1
NbAII.1	0.77±0.2	4.3±0.6	5.6±0.8
NbAIII.15	3.08±0.02	0.39±0.01	0.13±0.04
NbAIII.16	3.7±0.3	0.7±0.2	0.21±0.05

Error represent the standard error of two biological replicates.

Studies have shown that UCH37 changes its conformation, mainly at the C terminal ULD domain and the crossover loop upon binding with RPN13. We would like to validate the interactions between the nanobodies with the UCH37·RPN13$_{DEU}$ in addition to our earlier qualitative pull-down assay. Here, these nanobodies were incubated with UCH37·RPN13$_{DEU}$ at 1:1 ratio and were analyzed using size exclusion chromatography coupled with multi angle light scattering (SEC-MALS). SEC-MALS result indicated that NbAI.1 and its derivatives, NbAIII.15 and NbAIII.16 co-eluted with UCH37·RPN13$_{DEU}$ in a homogenous peak, with a measured mass at 53 ± 3 kDa, 59 ± 2 kDa, and 59.2 ± 0.5 kDa (increase from 45 ± 3 kDa for UCH37·RPN13$_{DEU}$). We did not observe co-eluted peak in the NbAII.1 and UCH37·RPN13$_{DEU}$ mixture (**Figure 3-10**). This supported our findings in the *in vitro* pull-down assay, suggesting that the binding interaction of NbAII.1 with UCH37 might be sensitive to the change of conformation in the UCH37·RPN13$_{DEU}$.

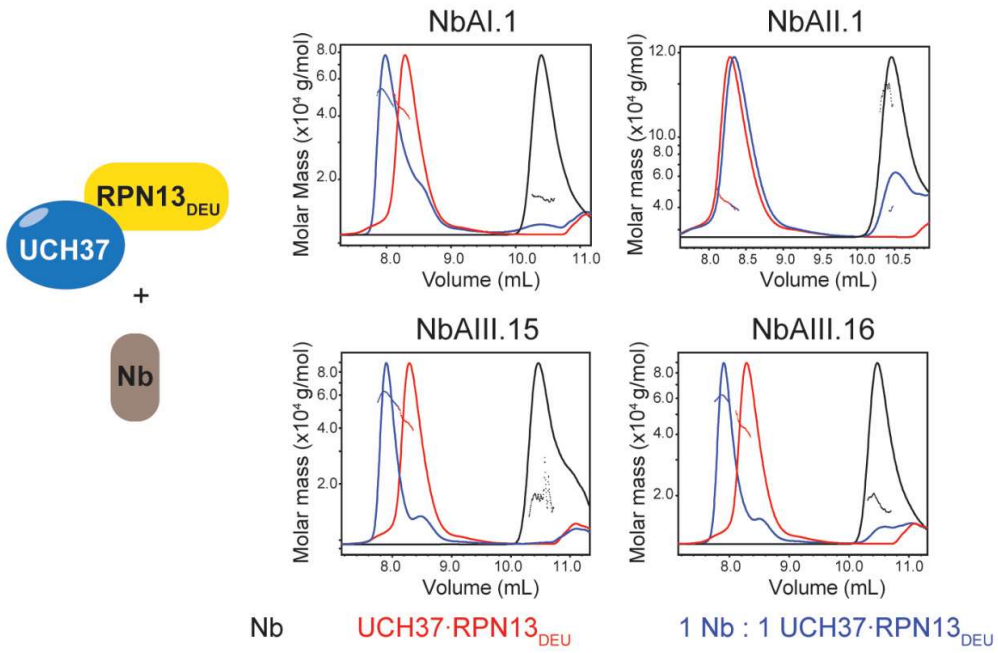

Figure 3-10: SEC-MALS analysis of nanobodies incubated with UCH37·RPN13$_{DEU}$.

Representative elution profiles of NbAI.1 (Top left), NbAII.1 (Top right), NbAIII.15 (Bottom left) and NbAIII.16(Bottom right) incubated with UCH37·RPN13$_{DEU}$ at 1:1 ratio (Blue). The black and red plots represent the elution profile of just nanobody and just UCH37·RPN13$_{DEU}$.

Together, our biophysical characterization suggested that NbAIII.15 could be the most promising nanobody to be developed as an intrabody for proteasome-associated UCH37 due to its high affinity towards full length UCH37 and its interaction with UCH37·RPN13. We are interested in examining the possible epitopes recognized by NbAIII.15. Limited proteolysis has been applied to probe the interacting sites of protein complexes. These binding surfaces would be occluded from the protease while the exposed area would be readily cleaved. These proteolytic fragments could be detected through mass spectrometry, such as matrix-assisted laser desorption/ionization- time of flight mass spectrometry

(MALDI-TOF MS) and be identified based on the specificity of protease. This would define the location of proteolysis and give insight into the binding interactions.[35] Here, we performed limited proteolysis on the native UCH37·RPN13$_{DEU}$ and UCH37·RPN13$_{DEU}$ in the presence of excess NbAIII.15 (in 1:2 ratio) using LysC. Our MALDI-TOF MS result elucidated that while the nanobody and RPN13$_{DEU}$ were mostly resistant to proteolysis, UCH37 was not; the MS spectrum also revealed that the apo UCH37·RPN13$_{DEU}$ has different cleavage pattern compare to NbAIII.15·UCH37·RPN13$_{DEU}$ complex (**Figure 3-11A**). The identification of these proteolytic fragments indicated that the catalytic domain of the UCH37 in the ternary complex remained mostly intact, in contrast to the apo UCH37·RPN13$_{DEU}$. In addition, certain regions on the UCH37 of the ternary complex were resistant towards proteolysis compare to the UCH37·RPN13$_{DEU}$. These regions, one is the solvent exposed surface delineated by K57 at β2, K206 and K210 at H7, the other located at the opposite end and included the K23 at H1, K228 at H8, and K281 at H9, where both H8 and H9 belongs to the ULD domain (**Figure 3-11B and Table 3-4**). These surfaces could serve as the possible binding site for NbAIII.15. Not surprisingly, both sites did not include the reported primary ubiquitin binding site, further verifying that NbAIII.15 is an exosite binder for UCH37·RPN13$_{DEU}$. While for NbAIII.15, we observed peptide fragments originated from the LysC proteolysis at sites away from the CDRs (**Table B-2**). Further characterization using X-ray crystallography and hydrogen deuterium exchange mass spectrometry are now being conducted

to obtain high resolution structure of the NbAIII.15·UCH37·RPN13$_{DEU}$ complex to give detailed structural information at atomic level.

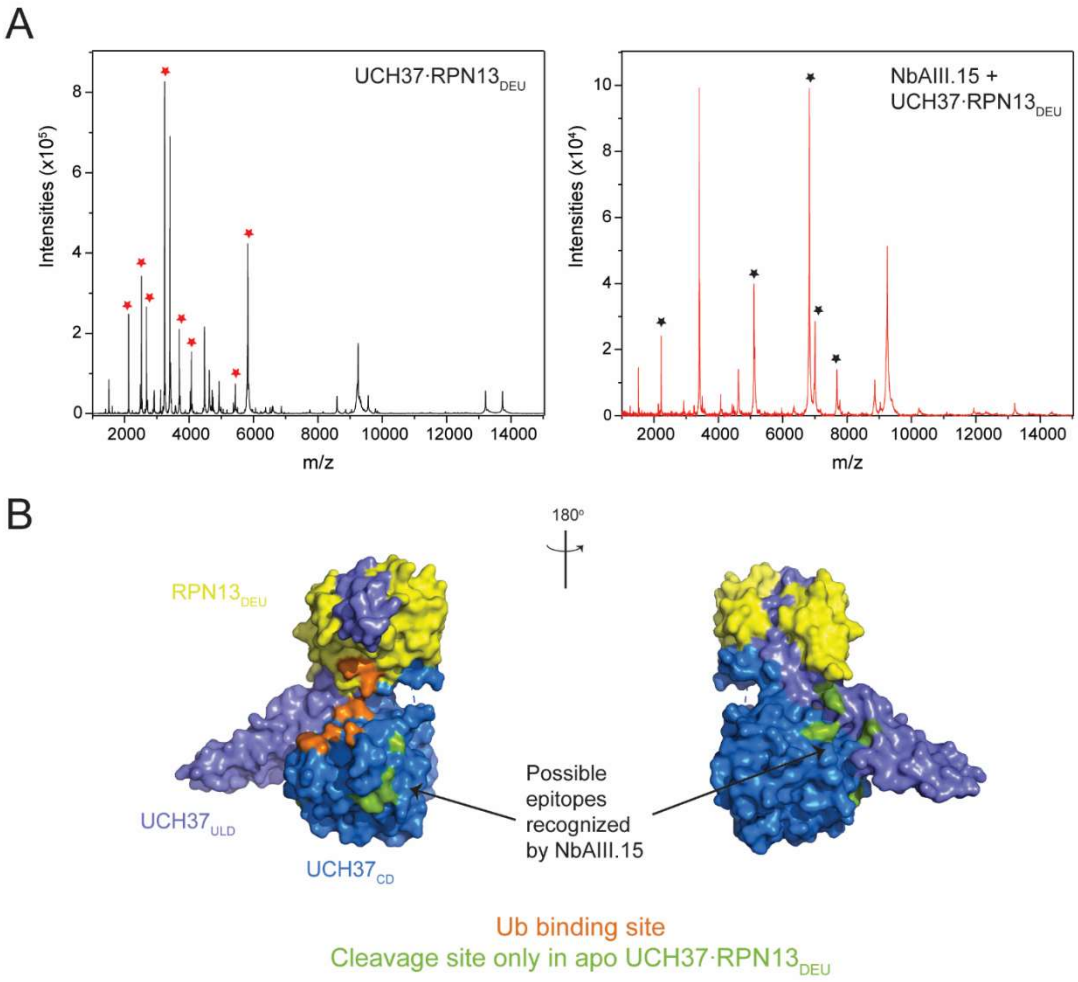

Figure 3-11: Limited proteolysis of UCH37·RPN13$_{DEU}$ and NbAIII.15 with UCH37·RPN13$_{DEU}$

(**A**) Representative MALDI-TOF MS spectrum for the limited proteolysis of UCH37·RPN13$_{DEU}$ (Left) and NbAIII.15 with UCH37·RPN13$_{DEU}$ (Right) by LysC. Red star indicated the proteolytic fragments only observed in UCH37·RPN13$_{DEU}$; black star showed the full length nanobody and proteolytic fragments originated from NbAIII.15. (**B**) Crystal structure of UCH37·RPN13$_{DEU}$ (PDB ID: 4UEL) with the cleavage sites observed only in UCH37·RPN13$_{DEU}$ highlighted. Cleavage sites which are unique to UCH37·RPN13$_{DEU}$ could potentially be the epitopes recognized by NbAIII.15.

Table 3-4: MALDI-TOF MS analysis of the limited proteolysis of UCH37·RPN13$_{DEU}$ and NbAIII.15 with UCH37·RPN13$_{DEU}$

UCH37·RPN13$_{DEU}$ Measured Mass $[M+H]^+$	NbAIII.15·UCH37·RPN13$_{DEU}$ Measured Mass $[M+H]^+$	Calculated Mass $[M+H]^+$	Observed amino acid residues
1516	1516	1519	289-300
2129	-	2130	211-228
2527	-	2527	24-45
2679	-	2679	58-81
3242	-	3241	127-154
3406	3405	3405	235-265
3699	-	3698	82-115
4074	-	4073	46-81
5436	-	5436	235-281
5823	-	5823	155-206
8864	8862	8864	159-234
9251	9250	9252	155-234

Only proteolysis fragments generated from UCH37 were listed here. These fragments were detected in two biological replicates. The identities of the proteolysis fragments were determined using ProteinProspector.[36]

Nanobodies as intrabodies for proteasome-associated UCH37

The nanobody that is suitable for live cell application should interact specifically with proteasomal UCH37 *in cellulo*. Here, we performed transient transfection of HEK293FT cells with HA-FLAG-HaloTag nanobodies.[37] Immunoprecipitation with HaloLink^{TM} resin followed by immunoblotting analysis showed that NbAI.1, NbAIII.15 and NbAIII.16 captured the endogenous UCH37. In addition, NbAIII.15 also immunoprecipitated with RPN13, other constitutive 19S RP subunits such as RPN11 and RPT2, and 20βs, which is a 20S CP subunit. While the negative control Nb, NbCtrl and NbAII.1 did not

immunoprecipitate with UCH37 (**Figure 3-12**). Finally, we performed immunoprecipitation followed by mass spectrometry (IP-MS) to investigate the specificity of NbAIII.15 in the context of cellular proteome and to identify *de novo* PPIs. Immunoprecipitated proteins were eluted using urea following by trypsin digestion and the resulting peptides were subjected to nanoLC-MS/MS. Our data showed that NbAIII.15 specifically captured UCH37 and the proteasome subunits compare to NbCtrl, further supporting our immunoblotting result (**Table 3-5 and Figure B-7**). In addition, it was suggested that NbAIII.15 interacts with INO80E, which is a subunit of INO80 complex. It has been reported that UCH37 could be recruited to INO80 complex to regulate the DNA transcription and DNA repair of the INO80 chromatin-remodeling complex through INO80G.[38] Structural studies revealed that UCH37 binds to INO80G at its DEUBAD domain, adopting an inhibitory conformation.[11,12] Nonetheless, the absence of other INO80 subunits may suggest that NbAIII.15 preferentially interacts with proteasome associated UCH37, which might be due to the unfavorable interactions with the UCH37 in a different conformation. In addition, INO80 complex exists in nucleus exclusively, and there is a possibility that the NbAIII.15 shows predominant cytoplasmic location, which lead to the lack of interaction with the INO80 complex associated UCH37. Further experiments need to be conducted to elucidate the interactions between the UCH37 in distinct conformations. Together, we have demonstrated that the isolated nanobodies, particularly NbAIII.15 could be developed as intrabody for proteasome-associated UCH37.

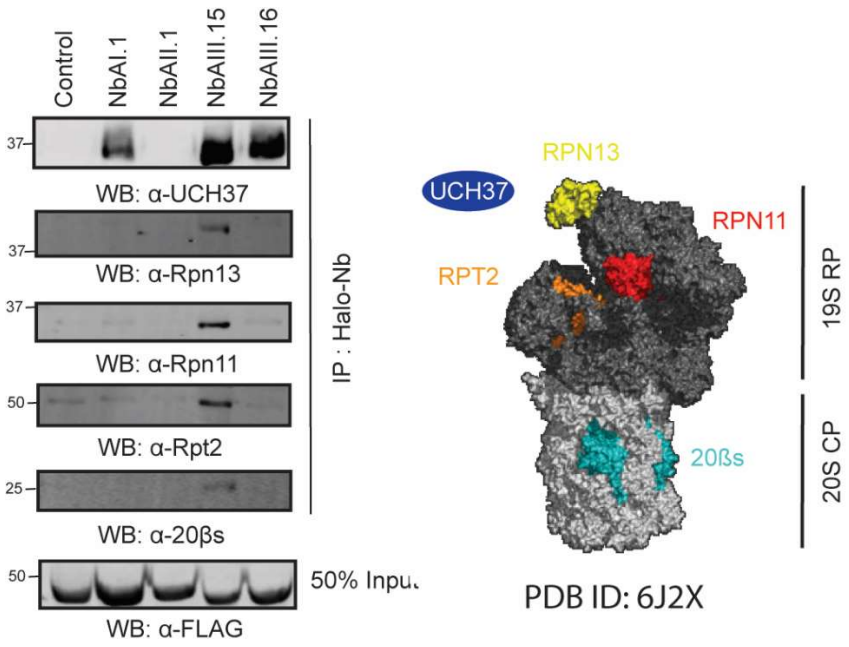

Figure 3-12: Immunoprecipitation of endogenous UCH37 by intracellularly expressed nanobodies

Representative Western blots showing the transient expression of nanobodies and the immunoprecipitation of UCH37, RPN13, RPN11, and RPT2, which are the 19S RP subunits and 20βs, which is the 20S CP subunits by intracellularly expressed NbAIII.15 (Left). A Cryo-EM structure of yeast proteasome (PDB ID: 6J2X) is shown as a representative scheme to depict the immunoprecipitated proteasomal subunits (Right). Immunoprecipitations were performed in two biological replicates.

Table 3-5: Complete list of NbAIII.15 interacting proteins

Accession	Description	Gene Symbol	# Unique peptides	
			Rep 1	Rep 2
Q9Y5K5	Ubiquitin carboxyl-terminal hydrolase isozyme L5	UCHL5	11	16
Q16186	Proteasomal ubiquitin receptor ADRM1	ADRM1	11	7
Q15008	26S proteasome non-ATPase regulatory subunit 6	PSMD6	30	17
P51665	26S proteasome non-ATPase regulatory subunit 7	PSMD7	16	11
O00231	26S proteasome non-ATPase regulatory subunit 11	PSMD11	30	17
Q9UNM6	26S proteasome non-ATPase regulatory subunit 13	PSMD13	26	15
O00487	26S proteasome non-ATPase regulatory subunit 14	PSMD14	15	10
P35998	26S proteasome regulatory subunit 7	PSMC2	33	24
O43242	26S proteasome non-ATPase regulatory subunit 3	PSMD3	35	26
P17066	Heat shock 70 kDa protein 6	HSPA6	8	3
P25685	DnaJ homolog subfamily B member 1	DNAJB1	5	4
Q8NBZ0	INO80 complex subunit E	INO80E	5	4
O95229	ZW10 interactor	ZWINT	3	6
Q96PM5	RING finger and CHY zinc finger domain-containing protein 1	RCHY1	6	3

Proteins which were identified in both biological replicates with $p<0.05$ are shown. Rows shaded in blue represent known proteasome subunits. Row shaded in beige shows a component of INO80 complex. UCH37 is reported to interact with INO80 complex through interaction with INO80G. MS data acquisition and analysis were performed by Yanfeng Li.

Discussion and Future directions

In this chapter, we have identified nanobodies specific towards proteasome associated UCH37 using a yeast surface display based synthetic nanobody library with an effective size of 2×10^7. The *in vitro* yeast display platform has allowed us to have a greater control on the antigen presented in our

screening strategy. In our effort to identify UCH37 binders, we chose ubiquitin conjugated UCH37·RPN13$_{DEU}$ as our antigen to: (i) prevent the oligomerization of UCH37 which lead to unfavorable conformation and (ii) to bias our screening result to identify exosite binders with higher specificity. To ensure the nanobodies could be developed as an intrabodies towards proteasome associated UCH37, we further performed affinity maturation using cell lysate generated from HEK293$^{RPN11\text{-}HTBH}$, which is a cell line used to prepare UCH37 containing proteasome, as our antigen. All the antigens that were used in our screening strategies are either heteromeric or unpurified, hence, negative selection using UCH37 KO HEK293$^{RPN11\text{-}HTBH}$ is incorporated in our screening strategies to eliminate the risk of enriching nanobodies targeting unwanted protein.

We have adopted the multiplexing in next generation sequencing to identify the enriched nanobodies using Illumina's MiSeq paired-end sequencing (2 x 250 bps), which has been the ideal sequencing platform to sequence the whole nanobody domain, which consist of around 400 bps. We further demonstrated that the depth of sequencing offered by next generation sequencing is crucial in our affinity maturation experiments. Our random mutagenized library consisted of 17% of unmutated template, which cause the discovery of positive clones through the low throughput sanger sequencing difficult. The improve sampling depth has assisted us in identifying the enriched mutations, which were later revealed to be beneficial. As the sequencing libraries were prepared through PCR, one of the limitations is that the bias during the PCR amplification may manifest artificial increase in the number of reads during

the sequencing and erroneous enrichment ratio. We attempted to correct the errors by the introducing degenerate codons at the 5' and 3' end of our libraries. These degenerate codons, acting as a unique identified code could decoupled the error due to bias of the PCR amplification from the enrichment from the screenings.[39] We have inspected our sequencing data with and without the unique identifiers, and our result showed that the same mutations, that are Q1R, R33P, H99Y, and A109P were enriched, indicating that the observed enrichment were genuine (data not shown). From our affinity maturation scheme, we observed the mutation hotspots at the CDRs which are at or close to the N terminal, suggesting the anchoring at the N terminal might influence the antigen recognition of our nanobodies. This prompt us to apply fusion protein at the N terminal of the identified nanobodies for downstream studies. This is not surprising in retrospect, as studies have indicated that the fusion protein at the N terminal of the displayed single chain variable fragment (scFv) and nanobody could influence the binding affinity towards the antigen.[40,41] Future synthetic library design could be designed to display the nanobodies on the N terminal of Aga2p.

We have discovered potential UCH37 binders targeting full length UCH37 with sub-nanomolar affinity. These nanobodies are capable to interact with UCH37, and UCH37 in complex with $RPN13_{DEU}$ or RPN13. Limited proteolysis of the tightest binder, NbAIII.15 in complex with UCH37·$RPN13_{DEU}$ suggested the NbAIII.15 binds at the catalytic domain of UCH37. Detailed structural investigation is currently in progress to elucidate the binding interactions of

NbAIII.15 and UCH37·RPN13$_{DEU}$ at atomic level. We also investigate whether these UCH37 nanobodies could be utilized as an intracellular tool to probe the regulation of proteasome-bound UCH37 and our result elucidated that these intracellularly expressed nanobodies, particular NbAIII.15 immunoprecipitates UCH37 containing proteasome. Further experiments are being conducted (i) to determine the biochemical effects of these nanobodies on the activity of UCH37 and (ii) to interrogate the compatibility of using the nanobodies *in cellulo* or *in vivo.*

Materials and Methods

Cell culture

For yeast cells used in the yeast surface display, referred to the cell culture section in the Material and Methods in Chapter 2.

HEK293 stably expressing RPN11-His$_6$-TEV-biotin tag-His$_6$ (HEK293$^{RPN11\text{-}HTBH}$), UCH37 KO HEK293$^{RPN11\text{-}HTBH}$, and HEK293FT were cultured in DMEM medium (Genesee Scientific) supplemented with 10% fetal calf serum (Genesee Scientific), 100 units/ml penicillin and 100 µg/ml streptomycin at 37^{o}C in a humified atmosphere with 5% CO_2.

Plasmid and Cloning

His$_8$-MBP-TEV-UCH37 *(Isoform 3)* in pVP16 vector was obtained from DNASU. GST-3C-RPN13$_{DEU}$ (or RPN13$_{253\text{-}407}$) was cloned from His$_{10}$-RPN13 in pET19b, which was a gift from Joan Conaway & Ronald Conaway (Addgene plasmid #

19423). 1436 pcDNA Flag HA was a gift from William Sellers (Addgene plasmids #10792). All PCR amplifications were performed with Phusion® High-Fidelity DNA polymerase (New England Biolabs). The primers used in this study were purchased from IDT DNA. All the sequences were verified through Sanger sequencing.

For bacterial expression of Avi-Ubiquitin$_{1-75}$ Intein, Avi-UCH37 and Avi-USP30 the Avi-tag were introduced through site-directed insertion. GST-TEV-UCH37 construct was generated by replacing the His$_8$-MBP fragment in the UCH37 in pVP16 vector with GST fragment (GST from pGEX6P1) through megaprimer PCR. For bacterial expression of the nanobody, the enriched nanobodies in the selections were cloned into pET28b (Novagen) at NdeI and XhoI (New England Biolabs) restriction enzyme sites. The point mutations were introduced into the nanobody through site-directed mutagenesis using megaprimer PCR. To generate HaloTag nanobody for mammalian expression, the GGGS$_3$ nanobody (in pCT vector) were PCR amplified and inserted into a HaloTag containing pET28b vector through megaprimer PCR. The HaloTag nanobodies cassetts were amplified and cloned into pCDNA3 HA FLAG (Addgene #10792) at BamHI and XhoI (New England Biolabs) restriction enzyme sites.

Protein purification

All protein purifications were performed at 4^{o}C.

Avi-Ubiquitin$_{1-75}$-propargylamine (AviUb-PA) was synthesized from Avi-Ub$_{1-75}$-intein. Briefly, Avi-Ub$_{1-75}$-intein was expressed in *E. coli* Rosetta (DE3) pLysS

using 2xYT media. The cells were grown at 37°C to OD 0.6 before induced with 0.5 mM IPTG at 18°C overnight. The cells were lysed in 20 mM sodium phosphate, pH 6.0, 500 mM NaCl, 1 mM EDTA. The Avi-Ub_{1-75}-intein was loaded onto chitin resin (New England Biolabs). The immobilized protein was incubated with lysis buffer in the presence of 100 mM sodium 2-mercaptoethanesulfonate (MESNA) (Acros Organic) for 48 hrs at 4°C to generate Avi-Ub_{1-75}-MESNA through on-resin cleavage. The Avi-Ub_{1-75}-MESNA (final concentration was 500 μM) was incubated in a buffered solution containing 20 mM Tris, pH 8.0, 10 mM *N*-hydroxysuccinimide, 1 M propargylamine (Alfa Aesar) at 37°C for overnight to yield AviUb-PA. All the Avi-Ub variants were verified using MALDI-TOF mass spectrometry (MS).

UCH37 was purified through glutathione affinity chromatography. In short, the GST-3C-UCH37 was expressed in *E. coli* BL21 (DE3). The cells were grown at 37°C to OD 0.6 before induced at 20°C with 0.4 mM IPTG for overnight. The harvested cells were lysed in 50 mM HEPES, pH 7.5, 150 mM NaCl and applied to glutathione (Genscript) resin. The elution buffer was 50 mM Tris, pH 8.0, 10 mM reduced glutathione. The UCH37 was further purified through anionic exchange chromatography using MonoQ on an ÄKTA purifier (GE Healthcare). The equilibrate buffer was 50 mM HEPES, pH 7.5, 0.5 mM TCEP and the elution buffer was 50 mM HEPES, pH 7.5, 1 M NaCl, 0.5 mM TCEP.

Avi-UCH37·$RPN13_{DEU}$ and UCH37·$RPN13_{DEU}$ complexes was purified through tandem affinity purification. In brief, His_8-MBP-TEV- Avi-UCH37/UCH37 and GST-3C-$RPN13_{DEU}$ were expressed in *E. coli* BL21 (DE3). The cells were grown

at 37°C to OD 0.6 before induced at 20°C with 0.4 mM IPTG for overnight. The harvested cells were mixed at 1:1 ratio and lysed in 50 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM imidazole. The complexes were purified using nickel (Gold Biotechnology) affinity chromatography to pull down His-tagged UCH37 and its binding partners. The complexes were eluted using 50 mM HEPES, 150 mM NaCl, 300 mM Imidazole, followed by the TEV cleavage reaction to remove the His_8-MBP tag. The resulting complexes were then buffer exchanged into 50 mM HEPES, pH 7.5, 150 mM NaCl and subjected to glutathione (Genscript) affinity chromatography to pull down GST-tagged $RPN13_{DEU}$ and its binding partners. The complexes were eluted using 50 mM Tris, pH 8.0, 10 mM reduced glutathione and 3C protease reaction was then carried out to remove the GST tag. The complexes were further purified through size exclusion chromatography using a HiLoad Superdex 200 pg on an ÄKTA purifier (GE Healthcare). The running buffer was 50 mM HEPES, pH 7.5, 50 mM NaCl, 0.5 mM TCEP. UCH37·RPN13 complex was purified through tandem affinity purification. In brief, both His_8-MBP-TEV-UCH37 and His_{10}-RPN13 were expressed in *E. coli* BL21 (DE3). The cells were grown at 37°C to OD 0.6 before induced at 20°C with 0.4 mM IPTG for overnight. The harvested cells were mixed at 1:1 ratio and lysed in 50 mM HEPES, pH 7.5, 200 mM NaCl, 1 mM EDTA and 1 mM TCEP. The complex was purified using amylose (New England Biolabs) affinity chromatography to pull down His_8-MBP-TEV-UCH37 and its binding partner. The complex was eluted with the lysis buffer in the presence of 10 mM maltose, followed by TEV cleavage reaction to remove His_8-MBP tag. The resulting

complexes were then buffer exchanged into 50 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM imidazole and subjected to nickel (Gold Biotechnology) affinity chromatography to pull down GST-tagged $RPN13_{DEU}$ and its binding partners. The complexes were eluted using into 50 mM HEPES, pH 7.5, 150 mM NaCl, 300 mM imidazole and was further purified through size exclusion chromatography using a HiLoad Superdex 200 pg on an ÄKTA purifier (GE Healthcare). The running buffer was 50 mM HEPES, pH 7.5, 50 mM NaCl, 0.5 mM TCEP.

The His_6-HALO-$GGGS_3$-nanobody/nanobody (Nb) variants were purified through immobilized metal affinity chromatography. The His_6-tagged Nbs were expressed in *E. coli* BL21 (DE3). The cells were grown at 37°C to OD ~0.6 before induced at 20°C with 0.4 mM IPTG for overnight. The harvested cells were lysed in 50 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole and purified using nickel (GoldBio) affinity chromatography. The elution buffer was 50 mM HEPES, pH 7.5, 300 mM NaCl, 300 mM imidazole. The Nbs were further purified through size exclusion chromatography using a HiLoad Superdex 75 pg on an ÄKTA purifier (GE Healthcare). The running buffer was 50 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM TCEP.

Generation of biotinylated protein

To prepare biotinylated Ub conjugated to UCH37·$RPN13_{DEU}$ (or $USP30_{67-517}$) , the Avi-Ub-PA was biotinylated through BirA reaction before conjugated to UCH37·$RPN13_{DEU}$. Briefly, the Avi-Ub-PA (final concentration was 40 μM) was

added to a solution containing 10 mM Tris, pH 8.0, 5 mM $MgCl_2$, 2 mM ATP, 150 mM D(+) biotin (Gold Biotechnology), and 0.4 μM BirA at 37°C until the biotinylation reaction went to completion (fresh BirA and ATP were provided every 30 mins). The reaction was verified using MALDI-TOF MS. The 5 μM biotinylated Ub-PA was incubated with 2.5 μM UCH37·$RPN13_{DEU}$ in a buffered solution containing 50 mM HEPES, pH 7.5, 50 mM NaCl, 1 mM TCEP at room temperature for 1 hr with gentle rocking. The biotinylated ternary complex was further purified through size exclusion chromatography using a HiLoad Superdex 200 pg on an ÄKTA purifier (GE Healthcare). The running buffer was 50 mM HEPES, pH 7.5, 50 mM NaCl, 1 mM TCEP.

To prepare biotinylated UCH37·$RPN13_{DEU}$ (or $USP30_{67-517}$) , the Avi-UCH37·$RPN13_{DEU}$ was biotinylated using the same BirA reaction conditions, except the reaction was performed at 4°C for 16 hrs with gentle rocking (fresh BirA and ATP were added into the reaction mixture in between). The biotinylated UCH37·$RPN13_{DEU}$ was further purified through size exclusion chromatography using a HiLoad Superdex 200 pg on an ÄKTA purifier (GE Healthcare). The running buffer was 50 mM HEPES, pH 7.5, 50 mM NaCl, 1 mM TCEP. Western blot was performed using streptavidin, Alexa Fluor™ 647 conjugate (ThermoFisher Scientific) at 1:15000 dilution to validate the biotinylation of UCH37.

Construction of the random mutagenized NbAl.1 library

The NbAl.1 gene was randomly mutagenized using Mutazyme II, a commercially available error prone DNA polymerase (Agilent Technologies) using pCT_Nb_amp_f and pCT_Nb_amp_r (**Table B-3**). The mutagenized NbAl.1 gene library pool was then amplified with the same primers using Phusion ® high fidelity DNA polymerase (New England Biolabs). To prepare electrocompetent yeast cells, a 50 ml culture of EBY100 cells were grown to OD 1.5 and treated with 25 mM DTT for 15 mins. The cells were then harvested, washed and resuspended in 2 x 150 µl of ice cold 10 mM Tris, pH 7.5, 270 mM sucrose, 2 mM $MgCl_2$. Cells were transformed separately using 1 µg of NheI and BamHI (New England Biolabs) of digested pCT vector containing NbAl.1 gene and 3 µg of amplified Nb gene library pool through electroporation. The combined transformed cells were recovered in YPD media before exchanged into SDCAA media. Recovered cells were serial diluted and plated in -Trp dropout media for 2-4 days at room temperature to estimate the size of the library.

Next generation sequencing (NGS) sample preparation and analysis

The sequencing library was generated through reduced cycle amplification to add Illumina adapters to the Nb in two sequential PCR. In the first PCR, the extracted plasmids were PCR amplified using Nbseq_deg_f and Nbseq_deg_r (**Table B-4**) to add in the partial Illumina adapters to the Nb. The PCR cycle conditions were: 98^{o}C for 30 s, followed by 25 cycles of 98^{o}C for 10 s, a temperature gradient of 64 – 68^{o}C for 30 s, 72^{o}C for 30 s, with a final extension of 72^{o}C for 5 mins. PCR

products was separated through agarose gel electrophoresis where the amplicons with the correct base pairs were excised and purified. The purified amplicons were PCR amplified again with reduced amount (0.05 μM) of Nbseq_N502/N503 and Nbseq_N701 (**Table B-4**) to add in the full-length Illumina adapters. The PCR conditions were 98^{o}C for 30 s, followed by 12 cycles of 98^{o}C for 10 s, 55^{o}C for 30 s, 72^{o}C for 60 s, with a final extension at 72^{o}C for 5 mins. All PCRs were performed using Phusion ® high fidelity DNA polymerase (New England Biolabs)

The quantity of Nb amplicon libraries was verified using Qubit dsDNA HighSensitivity Assay (Life Technologies). The quality of the libraries was checked using Agilent 2100 Bioanalyzer, and there was no detectable adapter dimer. The libraries were quantitated using NEBNext Library Quant Kit for Illumina (New England Biolabs) on Agilent Mx3005P qPCR instrument. The libraries were pooled with 40% Phi-X library spike-in (Illumina) and sequenced on Illumina MiSeq using Nano v2-500 cycle kit, with 251 bp paired-end chemistry. The quality of the sequence data was checked using FastQC.[42] and the average quality per read for all samples was in a range of a mean Phred score of 37.

For data analysis, the paired-end read were merged using FLASH.[43] The resulting reads were trimmed by BBDuk translated with SeqKit.[44,45] Downstream analysis was performed in R. The enrichment ratios of a given variants was calculated by dividing the normalized Nb frequencies enriched through sorting over the normalized Nb frequencies presented in the initial error prone NbAl.1 library.

Selection of Nanobodies (Nbs) against UCH37·RPN13$_{DEU}$

The display of the Nb on the yeast surface was induced by resuspending 10x diversity of the Nb library in the SGCAA (20 g/L galactose, 6.7 g/L yeast nitrogen base, 5 g/L casamino acids, 5.4 g/L Na_2HPO_4, 8.56 g/L $NaH_2PO_4 \cdot H_2O$) media supplemented with 100 units/ml penicillin and 100 µg/ml streptomycin at 20^oC. The induced cells were subjected to washing before and after incubation with PBSF buffer (0.1% (v/v) BSA in PBS buffer, pH 7.4). The induced cells were subjected to two subsequent magnetic activated cell sorting (MACS) and two fluorescent activated cell sorting (FACS) to select Nbs bind to Ub conjugated UCH37·RPN13$_{DEU}$. For MACS, the induced cells were negatively sorted against Dynabeads^{TM} M-280 Streptavidin (ThermoFisher Scientific) before positively sorted against Dynabeads coated with biotinylated Ub conjugated UCH37·RPN13$_{DEU}$. All MACS were performed at 4^oC for 1-2 hrs. For FACS, the induced cells were labeled with chicken anti c-Myc (Exalpha biologicals INC) at 1:250 dilution for 30 mins at room temperature. The labeled cells were then incubated with varied concentration of biotinylated Ub conjugated UCH37·RPN13$_{DEU}$ or biotinylated UCH37·RPN13$_{DEU}$ in complex with NbAl.1 for 30 mins at room temperature before simultaneously labeled with goat anti chicken IgY (H+L) secondary antibody, Alexa Fluor^{TM} 647 and streptavidin, Alexa Fluor^{TM} 488 conjugate (ThermoFisher Scientific) at 1:100 dilution for 15 mins on ice in a shielded box. Sortings were performed on a BD FACS ARIA II with 488 nm (530/30 band pass collection filter) and 640 nm (670/30 band pass filter) excitation lasers. At the end of FACS, yeast display plasmids were isolated from

the enriched clones using the Zymoprep yeast plasmid miniprep kit (Zymo Research). The enriched clones were identified through Sanger sequencing using pCT_Nb_amp_r (**Table B-3**).

Selection of Nanobodies against proteasome associated UCH37

The NbAl.1 library was cultured, induced, and washed as previously described. The induced yeast cells were first labeled with chicken anti c-Myc at 1:250 dilution for 30 mins at room temperature. This was followed by the incubation of the cells with clarified cell lysate of the HEK293$^{RPN11\text{-}HTBH}$ in a buffered solution containing 50 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 10% glycerol for 45 mins at room temperature. The cells were then similarly labeled with the secondary antibodies and sorted through FACS as mentioned in the *Selection of nanobodies against UCH37·RPN13$_{DEU}$*. For negative sorting, cell lysate of UCH37 KO HEK293T$^{RPN11\text{-}HTBH}$ was used as the antigen. The plasmids containing gene of the enriched clones was extracted through Zymoprep yeast plasmid miniprep kit and submitted for Sanger sequencing or used in the preparation of NGS sequencing library.

Validation of the enriched mutations

Yeast cells were transformed with the plasmid encoded for the nanobodies with enriched mutations (referred to Yeast growth and induction in the Materials and Methods in Chapter 2). The yeast cells were cultured, induced, and washed as previously described. For validation, the cells were incubated with the cell lysate

generated from HEK293$^{RPN11\text{-}HTBH}$ or UCH37 KO HEK293$^{RPN11\text{-}HTBH}$ in a buffered solution containing 50 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 10% glycerol for 45 mins, at room temperature with gentle rocking (or with biotinylated UCH37·RPN13$_{DEU}$ in PBS for 30 mins). The cells were then similarly labeled with primary and secondary antibodies as mentioned in the *Selection of nanobodies against proteasome associated UCH37* before subjected to flow cytometry analysis on a BD LSR Fortessa X20.

In vitro pull down assay

HaloTag nanobodies were incubated with the recombinantly expressed UCH37 or UCH37·RPN13 for 1 hr at room temperature with gentle rocking in a buffered solution of 50 mM HEPES, pH 7.5, 150 mM NaCl, 0.05% IGEPAL. The mixtures were then incubated with pre-equilibrated HaloLink^{TM} resin (Promega) for 30 mins at room temperature with gentle rocking. The immunoprecipitated proteins were extensively washed and eluted using SDS loading dye at 95^{o}C for 10 mins before subjected to SDS-PAGE analysis.

Surface plasmon resonance (SPR) measurement

The surface plasmon resonance experiments were performed using a BIACORE T200 (GE Healthcare) equipped with a Series S Sensor Chip NTA. The surfaces of the flow cells were conditioned using 350 mM EDTA for 1 min at a flow rate of 10 µl/min before injecting 0.5 mM $NiSO_4$ for 1 min at 10 µl/min. The purified His_6-Nb (15 kDa) was then injected for 2 mins at 10 µl/min and immobilized on the Ni-

NTA surface at a density of approximately 30 response units (RU). Single cycle kinetic titration experiments were performed by injecting purified UCH37 (37 kDa) at increasing concentration in 50 mM HEPES, pH 7.5, 50 mM NaCl, 50 μM EDTA, 0.005% (w/v) TWEEN 20 at a flow rate of 30 μl/min with a 300 s contact time and 900 s dissociation time. The flow cell surfaces were regenerated using 2 pulses of 350 mM EDTA for 2 min at a flow rate of 10 μl/min. The sensorgrams were double referenced before analyzed using BIAevaluation (GE Healthcare). The association and dissociation rate constant, k_a and k_d, were determined from the direct curve fitting to a 1:1 binding model; the dissociation equilibrium constant, K_D, was calculated from k_d/k_a. All experiments were performed at 25^{o}C.

Size exclusion chromatograph-multi angle light scattering (SEC-MALS) analysis

SEC-MALS was performed in a TSKgel G3000SW$_{XL}$ column (TOSOH Biosciences) on a Agilent 1260 Infinity HPLC system (Agilent Technologies) coupled to a MALS detector, DAWN HELEOS II (Wyatt Technology) and a refractive index detector, Optilab T-rEX (Wyatt Technology). For nanobody complexes, nanobodies were incubated with UCH37·RPN13$_{DEU}$ on ice for 10 mins before the analysis. The flow rate was 0.5 ml/ min and the mobile phase was 50 mM sodium phosphate, pH 6.5, 400 mM sodium perchlorate, 10% isopropanol. Data was analyzed using Astra (Wyatt Technology) and the molar masses of the complexes were determined by the Zimm fit model. All the experiments were performed at room temperature.

Limited proteolysis analysis

Limited proteolysis of UCH37·RPN13$_{DEU}$ and NbAIII.15 in complex with UCH37·RPN13$_{DEU}$ (around 4.5 µg at 1.1 µg/µl) were performed in 50 mM sodium phosphate, pH 8.0 with LysC at 1:40 g/g ratio for 37ºC for 15 mins. Here, NbAIII.15 was incubated with UCH37·RPN13$_{DEU}$ on ice at 2:1 ratio for 10 mins before digestion. For quenching, 1 µl of the reaction mixture was added into 10 µl of saturated α-Cyano-4-hydroxycinnamic acid (CHCA) (Sigma-Aldrich) in 30% acetonitrile, 0.1% TFA and 1 µl of the quenched sample were spotted on a 96-ground steel MALDI target plate. Data was acquired on a Bruker MALDI-TOF MicroFlex or UltrafleXtreme in linear positive ion mode with 20% laser power and 500 shots were collected per spectrum. Data processing was performed using FlexAnalysis software (Bruker). The identities of the proteolysis fragments were determined using ProteinProspector.

Transient transfection

The vectors used in the transient transfection were generated by the cloning of the 1436 pCDNA3 Flag HA (1436 pcDNA Flag HA was a gift from William Sellers (Addgene plasmids #10792)). Transient transfections of mammalian cells were performed using Xfect^{TM} transfection reagent (Takara Bio) according to the manufacturer's protocol and scaling recommendation. Briefly, HEK293FT were grown to 35% confluency. Plasmid DNA was incubated with the Xfect polymer at room temperature for at least 10 mins before added drop wise to the adhered cells. The cells were incubated with the complete growth media containing DNA

nanoparticles for 4 hrs before replacing it with fresh media and further incubated for 48 hrs and the cells were prepared for downstream analysis.

Immunoprecipitation for western blotting

10 cm^2 dishes of HEK293FT cells were grown and transfected with respective HaloTag Nb as described above. The harvested cells were lysed in 50 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 1 mM DTT, 10% glycerol before incubated with HaloLink™ Resin (Promega Corporation) for 30 mins at room temperature. The immunoprecipitated proteins was washed extensively with 50 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 1 mM DTT, 0.05% IGEPAL prior to elution using SDS loading dye at 95^oC for 10 mins. Western blot analysis were performed using the following antibodies: rabbit anti-UCH37 [EPR4896] (abcam), rabbit anti-ADRM1 (D9Z1U) (Cell Signaling Technology), rabbit anti-PSMD14 [EPR4258] (abcam), rabbit anti-PSMC1 (ab3317) (abcam) and mouse anti-PSMB7 (R&D systems) at 1:1000 dilution. IRDye® 680RD goat anti-rabbit and IRDye® 800CW goat anti-mouse (LI-COR Biosciences) were used as the secondary antibodies at 1:15000 dilution.

Immunoprecipitation for mass-spectrometry

15 cm^2 dishes of HEK293FT cells were grown and transfected with respective HaloTag nanobodies as described above. The harvested cells were lysed and immunoprecipitated similarly as described above. The immunoprecipitated proteins was washed extensively with 50 mM HEPES, pH 7.5, 150 mM NaCl, 5

mM $MgCl_2$, 2 mM ATP, 1 mM DTT before eluted with 8M urea in 25 mM NH_4CO_3 for 1 hr at room temperature. The eluted proteins were reduced with 10 mM DTT at 37°C for 1 hr and alkylated using 25 mM iodoacetamide (Sigma-Aldrich) in 25 mM NH_4CO_3 at room temperature for 1 hr in the dark before subjected to in solution trypsin digest using sequencing grade modified trypsin (Promega Corporation) for overnight at 37°C. Tryptic digests were desalted with Sep-Pak C18 cartridge (Waters) and separated by a homemade fused silica capillary column (75 µm x 150 mm) packed with C-18 resin (120A, 1.9 µm, Dr. Maisch HPLC GmbH) with the EASY-nLC 1000 nano-HPLC system (ThermoFisher Scientific), which was coupled with the Orbitrap Fusion Tribrid mass spectrometer. Peptides were eluted by applying a 90 min gradient elution at 300 nl/min: 5% to 40% buffer B over 40 min, 40% to 60% buffer B over 20 min, and 60-95% buffer B over 10 min (solvent A: 0.1% formic acid (FA) in water, solvent B: 0.1% FA in ACN). Full MS scans were acquired at a resolution of 120000 between 350-2000 m/z. Collision-induced dissociation (CID) was induced on the twelve most abundant ions per full MS scan using an isolation width of 20 ppm. Fragmented precursor ions were allowed one repeated MS/MS analysis and the excluded for 15 s.

LC-MS/MS Data were searched against human proteome database in UniProt by using the Proteome Discoverer 2.4 with the SEQUEST HT search engine. The search parameters were: peptides mass tolerance of 10 ppm; MS/MS tolerance of 0.6 Da; two missed cleavages allowed; Carbamidomethylation of cysteine was a fixed modification. Oxidation of methionine and carbamylation of lysine,

arginine and peptide N-terminus were used as variable modifications. Label-free quantitation was performed using the Minora algorithm of Proteome Discoverer 2.4. The following parameters were used: Protein abundances were based on peptide peak area, top-3 peptides will be used for quantification. Protein abundance ratios were based on pairwise ratios. Student's t-test will be used and differences were considered as significant at the $p < 0.05$ level.

References

1. Voges, D., Zwickl, P. & Baumeister, W. The 26S Proteasome: A molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* **68**, 1015–1068 (1999).
2. Grice, G. L. & Nathan, J. A. The recognition of ubiquitinated proteins by the proteasome. *Cell. Mol. Life Sci.* **73**, 3497–3506 (2016).
3. de Poot, S. A. H., Tian, G. & Finley, D. Meddling with Fate: The Proteasomal Deubiquitinating Enzymes. *J. Mol. Biol.* **429**, 3525–3545 (2017).
4. Lander, G. C. *et al.* Complete subunit architecture of the proteasome regulatory particle. *Nature* **482**, 186–191 (2012).
5. Tanaka, K. The proteasome: Overview of structure and functions. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* **85**, 12–36 (2009).
6. Qiu, X. B. *et al.* hRpn13/ADRM1/GP110 is a novel proteasome subunit that binds the deubiquitinating enzyme, UCH37. *EMBO J.* **25**, 5742–5753 (2006).
7. Hamazaki, J. *et al.* A novel proteasome interacting protein recruits the deubiquitinating enzyme UCH37 to 26S proteasomes. *EMBO J.* **25**, 4524–4536 (2006).
8. Yao, T. *et al.* Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1. *Nat. Cell Biol.* **8**, 994–1002 (2006).
9. Nishio, K. *et al.* Crystal structure of the de-ubiquitinating enzyme UCH37 (human UCH-L5) catalytic domain. *Biochem. Biophys. Res. Commun.* **390**, 855–860 (2009).

10. Burgie, S. E., Bingman, C. A., Soni, A. B. & Phillips, G. N. Structural characterization of human Uch37. *Proteins Struct. Funct. Bioinforma.* **80**, 649–654 (2012).

11. Sahtoe, D. D. *et al.* Mechanism of UCH-L5 Activation and Inhibition by DEUBAD Domains in RPN13 and INO80G. *Mol. Cell* **57**, 887–900 (2015).

12. VanderLinden, R. T. *et al.* Structural Basis for the Activation and Inhibition of the UCH37 Deubiquitylase. *Mol. Cell* **57**, 901–911 (2015).

13. D'Arcy, P. *et al.* Inhibition of proteasome deubiquitinating activity as a new cancer therapy. *Nat. Med.* **17**, 1636–1640 (2011).

14. Lu, X. *et al.* Structure of the Rpn13-Rpn2 complex provides insights for Rpn13 and Uch37 as anticancer targets. *Nat. Commun.* **8**, 15540 (2017).

15. Aufderheide, A. *et al.* Structural characterization of the interaction of Ubp6 with the 26S proteasome. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8626–8631 (2015).

16. Stocks, M. Intrabodies as drug discovery tools and therapeutics. *Curr. Opin. Chem. Biol.* **9**, 359–365 (2005).

17. Gupta, A. *et al.* Facile target validation in an animal model with intracellularly expressed monobodies. *Nat. Chem. Biol.* 1–6 (2018) doi:10.1038/s41589-018-0099-z.

18. Schumacher, D., Helma, J., Schneider, A. F. L., Leonhardt, H. & Hackenberger, C. P. R. Nanobodies : Chemical Functionalization Strategies and Intracellular Applications Angewandte. 2314–2333 doi:10.1002/anie.201708459.

19. Dmitriev, O. Y., Lutsenko, S. & Muyldermans, S. Nanobodies as probes for protein dynamics in vitro and in cells. *J. Biol. Chem.* **291**, 3767–3775 (2016).

20. Mcmahon, C. *et al.* Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, (2018).

21. Zimmermann, I. *et al.* Synthetic single domain antibodies for the conformational trapping of membrane proteins. *Elife* **3**, 1–32 (2018).

22. Jiao, L. *et al.* Mechanism of the Rpn13-induced activation of Uch37. *Protein Cell* **5**, 616–630 (2014).

23. Mevissen, T. E. T. & Komander, D. Mechanisms of Deubiquitinase Specificity and Regulation. *Annu. Rev. Biochem.* **86**, 159–192 (2017).

24. Komander, D., Clague, M. J. & Urbé, S. Breaking the chains: Structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**, 550–563 (2009).

25. Gautier, A. & Hinner, M. J. Site-specific biotinylation of purified proteins using BirA. *Site-Specific Protein Labeling Methods Protoc.* 171–184 (2015) doi:10.1007/978-1-4939-2272-7.

26. Ackerman, M., Levary, D., Tobon, G., Hackel, B. & Orcutt, K. D. Highly Avid Magnetic Bead Capture: An Efficient Selection Method. *Artif. Cells Blood Substitutes Immobil. Biotechnol. Cells Blood Substit Immobi* **25**, 774–783 (2010).

27. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. **1**, 755–769 (2007).

28. Hunter, S. A. & Cochran, J. R. *Cell-Binding Assays for Determining the Affinity of Protein–Protein Interactions: Technologies and Considerations*. *Methods in Enzymology* vol. 580 (Elsevier Inc., 2016).

29. Wang, X. *et al.* Mass Spectrometric Characterization of the Affinity-Purified Human 26S Proteasome Complex †. (2007) doi:10.1021/bi061994u.

30. Kuo, C.-L., Collins, G. A. & Goldberg, A. L. Methods to Rapidly Prepare Mammalian 26S Proteasomes for Biochemical Analysis. *Methods Mol. Biol.* 277–288 (2018).

31. Tillotson, B. J., De Larrinoa, I. F., Skinner, C. A., Klavas, D. M. & Shusta, E. V. Antibody affinity maturation using yeast display with detergent-solubilized membrane proteins as antigen sources. *Protein Eng. Des. Sel.* **26**, 101–112 (2013).

32. Mason, D. M. *et al.* High-throughput antibody engineering in mammalian cells by CRISPR/Cas9-mediated homology-directed mutagenesis. *Nucleic Acids Res.* **46**, 7436–7449 (2018).

33. Deshpande, I. *et al.* Smoothened stimulation by membrane sterols drives Hedgehog pathway activity. *Nature* **571**, 284–288 (2019).

34. Karlsson, R., Katsamba, P. S., Nordin, H., Pol, E. & Myszka, D. G. Analyzing a kinetic titration series using affinity biosensors. *Anal. Biochem.* **349**, 136–147 (2006).

35. Suh, M. J., Pourshahian, S. & Limbach, P. A. Developing Limited Proteolysis and Mass Spectrometry for the Characterization of Ribosome Topography. *J. Am. Soc. Mass Spectrom.* **18**, 1304–1317 (2007).

36. ProteinProspector. http://prospector.ucsf.edu/prospector/mshome.htm.

37. Los, G. V. *et al.* HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008).

38. Yao, T. *et al.* Distinct Modes of Regulation of the Uch37 Deubiquitinating Enzyme in the Proteasome and in the Ino80 Chromatin-Remodeling Complex. *Mol. Cell* **31**, 909–917 (2008).

39. Glanville, J. *et al.* Deep sequencing in library selection projects: What insight does it bring? *Curr. Opin. Struct. Biol.* **33**, 146–160 (2015).

40. Uchański, T. *et al.* An improved yeast surface display platform for the screening of nanobody immune libraries. *Sci. Rep.* **9**, 1–12 (2019).

41. Wang, Z., Mathias, A., Stavrou, S. & Neville, D. M. A new yeast display vector permitting free scFv amino termini can augment ligand binding affinities. *Protein Eng. Des. Sel.* **18**, 337–343 (2005).

42. Andrews, S. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ doi:10.1016/0038-0717(73)90093-X.

43. Magoč, T. & Salzberg, S. L. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).

44. BBMap download | SourceForge.net. https://sourceforge.net/projects/bbmap/.

45. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**, (2016).

CHAPTER 4

DEVELOP AN INTRACELLULAR PROBE FOR PROTEASOME-ASSOCIATED UCH37

Abstract

As a main component in the ubiquitin proteasome system, the proteasome is subjected to multilayer regulations. One of the possible regulations is the recruitment of proteasomal subunits. UCH37 is a proteasome associated-deubiquitinase which is sub-stoichiometry to and interact dynamically with the proteasome. Here, we have engineered a nanobody as a probe for intracellular imaging of proteasome-bound UCH37. The nanobody does not affect the enzymatic activity of the proteasomal UCH37, which enable the detection the enzyme inside living cells in real time, with minimal perturbation towards the proteasome-associated UCH37. Furthermore, we improved the sensitivity of the probe by converting the nanobody to a conditionally stable variants to visualize proteasome-bound UCH37 close to endogenous level.

Introduction

26S proteasome is an important component in the ubiquitin dependent proteolysis, which is responsible for around 80% of the protein degradation in eukaryotes.[1] In the ubiquitin proteasome system (UPS), the substrates to be degraded are ubiquitinated, mainly with Lys48- linked ubiquitin chains, which would triggered the recruitment of 26S proteasome. The 26S proteasome is a macromolecular complex, consists of around 30 proteins subunits and could be categorized into two components: the 19S regulatory particle (19S RP) and 20S core particle (20S CP). The 19S RP contains the ubiquitin protein conjugates recognition or editing elements and an AAA+ ATPase which unfold the substrates. The ATPases also translocate the unfolded substrates through a gated channel to the barrel-shaped 20S CP with a proteolytic core for proteolysis.[2] The ubiquitin proteasome system involves in a plethora biological events, such as cell cycle, protein qualify control and signal transduction.

To ensure efficient ubiquitin-dependent proteolysis, the proteasome is tightly regulated at multi-level, from the expression level of the proteasome subunits, post translational modifications and the assembly of the proteasome complex.[3] Nonetheless, the recruitment of the 19S RP subunits which are responsible for the ubiquitin protein conjugates recognition and deubiquitination are still poorly understood. It has been well established that the binding of the ubiquitin protein conjugate does not necessary lead to substrate degradation by proteasome.[4] Hence, understanding the role of independent ubiquitin interacting proteasome associated subunits and how they contribute to proteasomal

degradation are crucial. First, why there are different ubiquitin receptors and deubiquitinases on the 19S RP, if they are performing the similar tasks? To what extent there is functional redundancy across these ubiquitin interacting proteins? There are two main proteasomal ubiquitin receptors on the proteasome: RPN10 and RPN13. Study has demonstrated that while RPN10 is the main ubiquitin receptors for ubiquitinated substrates, RPN13 might show preference towards the UBL domain of the ubiquitin shuttling factors.[5] This finding suggests that the presence of multiple ubiquitin receptors on the 19S RP would allow the versatile detection of ubiquitinated substrates with different topologies by the proteasome. Nonetheless, the deletion of either RPN10 or RPN13 only result in mild impairment, suggesting the role of RPN10 and RPN13 maybe redundant *in vivo*.[6]

Similarly, there are several proteasome-associated deubiquitinases located at the 19S RP; they are USP14, UCH37 and RPN11. Both USP14 and UCH37 are cysteine protease while RPN11 is a zinc metalloprotease.[7] Further biochemical studies have elucidated that both USP14 and RPN11 break down the ubiquitin chain *en bloc*, with RPN11 mainly hydrolyzes the isopeptide bond between the proximal ubiquitin and the substrate.[8] On the other hand, UCH37 is reported to remove ubiquitin from the distal tips and is shown to preferentially hydrolyzed branched ubiquitin chains.[9]

Intriguingly, quantitative mass spectrometry analyses has suggested that in contrast to RPN11, which is a constitutive proteasomal subunit, both USP14 and UCH37 exist at a sub-stoichiometric level, and interact with the 19S RP dynamically.[10,11] In addition, these deubiquitinases are reported to exist in

proteasome-free state.[12] It has been reported that the activity of the proteasome could be regulated by the assembly of proteasome with different subunits.[13,14,15] Are the labile interactions of USP14 or UCH37 with proteasome one of the regulation mechanism in the ubiquitin proteasome system? What is the recruitment pattern of USP14 and UCH37 on the proteasome inside a living cell? What are the triggers for the dynamic (dis)incorporation of USP14 and UCH37 on the proteasome?

Physiologically, the knockdown of RPN11 inhibits the degradation of cellular protein, which clearly delineate the role of RPN11 in the ubiquitin dependent proteolysis.[12,16] On the other hand, the knockdown or either USP14 or UCH37 has limited effect on the degradation rate, but the knockdown of both USP14 and UCH37 inhibit the protein degradation.[12] Does this suggest there are certain degree of functional redundancy between USP14 and UCH37? However, studies has elucidated that cells transiently express inactive USP14 and UCH37 lead to accumulation of distinct ubiquitinated substrates, suggesting these deubiquitinases might be substrates selective.[17]

We have discovered a nanobody specific towards proteasome-associated UCH37 and we hypothesize that the nanobody could be utilized as an intracellular probe to tackle the unknowns surrounding the proteasomal UCH37 as described above. For example, the nanobody could be fused to a fluorescent tag to visualize the localization of UCH37 on the proteasome and give insights into the dynamics of proteasome bound UCH37 under different biological events, such as stress or during cell cycle using fluorescent microscopy (**Figure 4-**

1A).[18,19] Furthermore, we could engineer a nanobody with enzyme-catalyzed proximity labeling, such as TurboID to study the substrate of UCH37 (**Figure 4-1B**).[20,21] UCH37 is a potential therapeutic target in cancer; a nanobody which inhibit the catalytic activity of UCH37 could be developed to investigate the physiological role of UCH37 in a disease model (**Figure 4-1C**).[22,23]

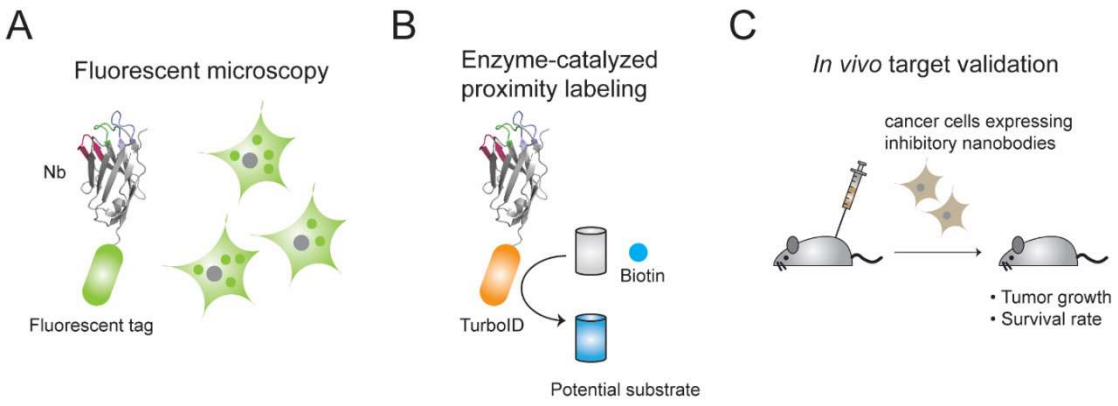

Figure 4-1: Potential applications of nanobodies.

Scheme showing the potential applications of nanobody to (**A**) visualize the recruitment of proteasome bound UCH37, (**B**) identify the potential substrates of UCH37, and (**C**) validate the functional role in a disease model.

Here, we showed that our engineered nanobodies does not interfere with the enzymatic activities of proteasomal UCH37. We further converted the nanobody to a conditionally stable nanobody under an inducible promoter; by removing the unbound nanobody, we reduce the background signal and permit the visualization of the proteasome bound UCH37 close to endogenous level.[24]

Result

UCH37 nanobodies do not inhibit proteasome-associated UCH37

In the previous chapter, we have discovered two distinct UCH37 nanobodies, which are NbAI.1 and NbAII.1 using ubiquitin conjugated UCH37·RPN13$_{DEU}$ at the catalytic Cys and UCH37·RPN13$_{DEU}$ as antigen. How do the different nanobodies affect the activities of UCH37? Our group has demonstrated that UCH37 exclusively hydrolyzes Lys48 linked branched ubiquitin chains.[9] UCH37 debranching of a high molecular weight ubiquitin chains consisting of Lys6 and Lys48 branched linkages (K6/K48 HMW ubiquitin chain) is followed through SDS-PAGE analysis (**Figure 4-2A and B**). Our result showed that both nanobodies effectively inhibit the debranching activities of UCH37, with an IC_{50} in the low micromolar range, which is the concentration of UCH37 used in the assay (**Figure 4-2C and Figure B-8**).

We further investigated the inhibitory effect of NbAI.1 and NbAII.1 on the debranching activity of UCH37·RPN13. We observed that NbAI.1 effectively inhibits the activity of UCH37·RPN13, while NbAII.1 did not (**Figure 4-2D and Table 4-1 and Figure B-9**). It is not surprising, as we had demonstrated that NbAII.1 does not complex with UCH37·RPN13$_{DEU}$ productively. Studies have shown that free UCH37 and UCH37·RPN13 adopt different conformation; our result indicated that while NbAII.1 might bind to free UCH37, it does not recognize or has low affinity towards UCH37·RPN13 where UCH37 has a different conformation.[25,26] At the highest concentration in the assay, which concentration was 20 times in excess to the concentration of UCH37·RPN13,

NbAII.1 only achieved 85% of inhibition (**Figure B-10**). We proceeded to study the effect of the derivatives of NbAI.1, which are NbAIII.15 and NbAIII.16 towards the debranching activities of UCH37·RPN13. These NbAI.1 derivatives which are identified using non purified UCH37 containing proteasome as antigen, also inhibit UCH37·RPN13 within the same order of magnitude with NbAI.1 (**Figure 4-2D and Table 4-1 and Figure B-9).**

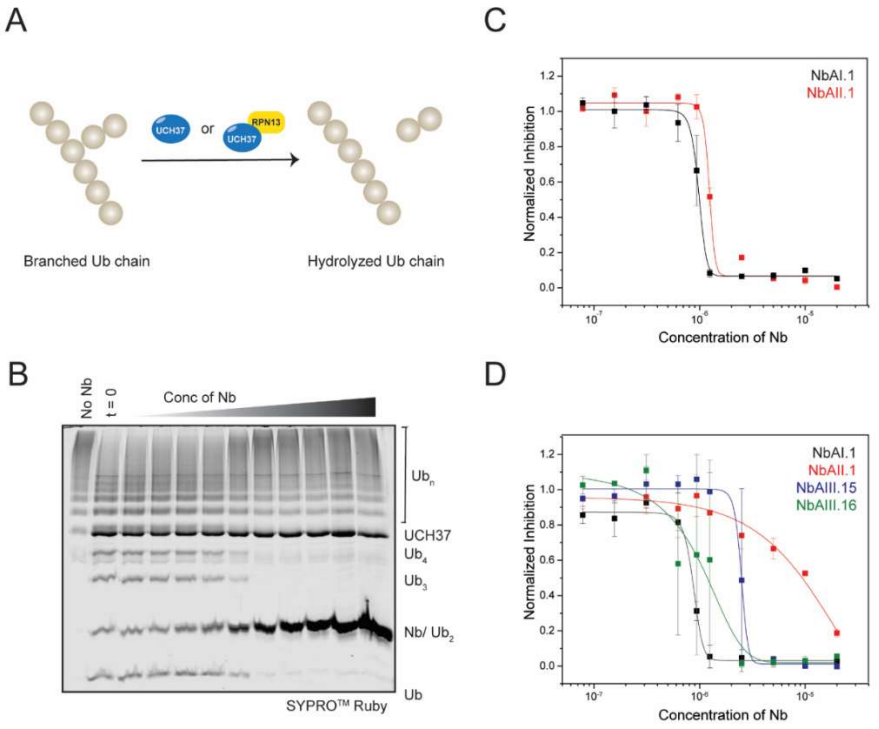

Figure 4-2: Investigate the effect of nanobodies on the activities of UCH37 and UCH37·RPN13

(**A**) Scheme of the UCH37 or UCH37·RPN13 cleavage assay using branched ubiquitin chains. (**B**) Representative gel for the cleavage assay using high molecular weight Lys6-/ Lys48- linked branched ubiquitin chains by UCH37 against a concentration gradient of NbAI.1. (**C**) Normalized inhibition of the debranching activities of UCH37 treated with NbAI.1 and NbAII.1. Error bars represent the standard error of two biological replicates. (**D**) Normalized inhibition of the debranching activities of UCH37·RPN13 treated with NbAI.1, NbAII.1, NbAIII.15 and NbAIII.16. Error bars represent the standard error of two biological replicates.

Table 4-1: IC_{50} of nanobodies towards UCH37·RPN13

Nb	IC_{50} (uM)
AI.1	0.94±0.09
AII.1	8±1
AIII.15	2.5±0.7
AIII.16	1±4

Next, we would like to check whether our nanobodies would inhibit the debranching activities of proteasome-associated UCH37. Using a HEK293 cell line stably expressing RPN11-His_6-TEV-Biotin-His_6 ($HEK293^{RPN11\text{-}HTBH}$), we have purified USP14 depleted proteasome through affinity purification, hence, the only proteasome-associated deubiquitinase on the purified proteasome that is capable of cleaving ubiquitin chains is UCH37 (**Figure B-11**). In parallel, we prepared biotinylated HMW K6/K48 ubiquitin chains with spike-in biotinylated ubiquitin to improve the sensitivity of the assay. Astonishingly, all of our nanobodies do not show significant inhibition towards proteasome-bound UCH37 in hydrolyzing HMW K6/K48 branched ubiquitin chains (**Figure 4-3A**). Even with NbAIII.15, which is demonstrated to interact with proteasome-associated UCH37, we only observed partial inhibition at high concentration, which is at 1000 times in excess of proteasomal UCH37 (**Figure 4-3B**). Further inhibition kinetics assays using UCH37·RPN13 and proteasomal UCH37 are crucial to understand the observed discrepancy in the inhibitory effect of our nanobodies towards UCH37 in different stages.

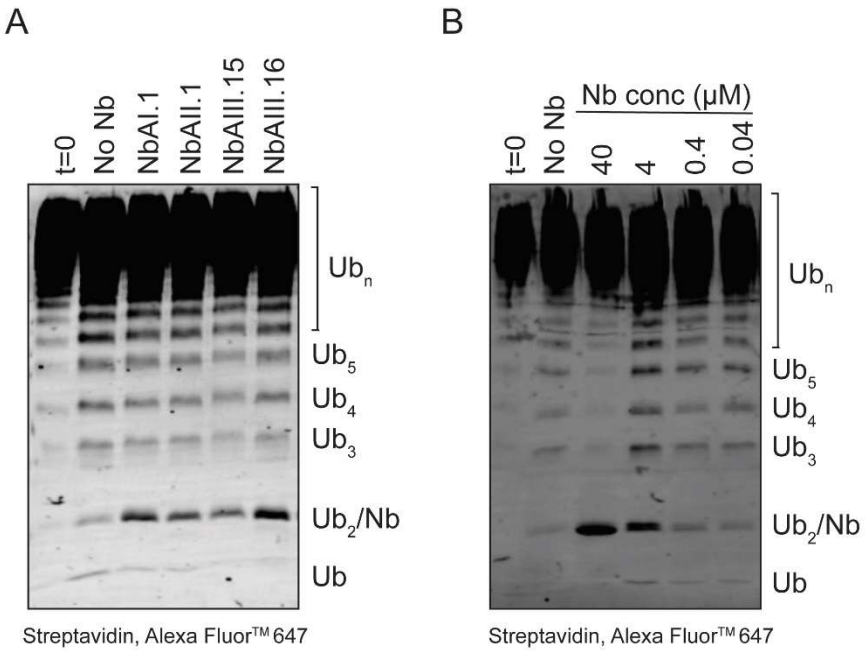

Figure 4-3: Branched polyubiquitin cleavage assay by proteasome associated UCH37

Representative Western blot for the cleavage assay using biotinylated high molecular weight Lys6- /Lys48- linked branched ubiquitin chain by proteasome associated UCH37 (**A**) treated with NbAI.1, NbAII.1, NbAIII.15 and NbAIII.16 and (**B**) treated with NbAIII.15 against a concentration gradient.

Develop conditionally stable proteasome-associated UCH37 nanobody

Our result indicated that the nanobodies bound proteasome associated UCH37 retains its enzymatic activity *in vitro*. We decided to develop our nanobodies as an intracellular imaging probe to study the regulation of the proteasome-associated UCH37. Here, we focused on NbAIII.15, which has been demonstrated to interacts with UCH37 containing proteasome in a native cellular context through immunoprecipitation. One of the major challenges of using the genetically expressed nanobodies is that their stabilities are independent of their antigens. As a result, it has been difficult to monitor the dynamic of the

nonmembrane associated protein which is ubiquitous in nucleus and cytoplasm. Nonetheless, there are studies reported the level of certain fluorescently labeled nanobodies, or chromobodies, are influenced by the expression level of their antigen, suggesting possible antigen-mediated stabilization effect by chromobodies. Further sensitizing the stability of the nanobody to the antigen level would allow the tracking of the dynamics endogenous proteins under different biological events, such as during cell cycles or upon treatment. In addition to accelerate the protein turnover by introducing certain amino acid residues to induce proteasomal degradation based on N-end rule, studies has shown that certain mutations could be readily transferred on the framework of the nanobodies to cause the nanobodies to be conditionally stable.[19] Those mutations are S70R, C92Y, Q105H, and S113F (according to the Chothia numbering) would destabilize the nanobodies in the absence of their antigens.[24,27,28] It was also demonstrated that these conditionally stable nanobodies are being degraded through the ubiquitin proteasome system.

We would like to check whether the nanobodies destabilization technique is applicable for NbAIII.15. Here, we prepared NbAIII.15 variants containing different combinations of the destabilizing mutations (S70R, C90Y, Q113H and S121F according to the amino acid sequences of NbAIII.15) (**Figure 4-4A**). The stabilities of the NbAIII.15 variants in HEK293FT and UCH37 KO HEK293FT cells were accessed semi-quantitatively through Western blot (**Figure 4-4B and Figure B-12**). Overall, we observed that the transiently expressed NbAIII.15 variants have decrease expression level compare to their wild type counterpart in

both WT and UCH37 KO HEK293FT cells, ranging from 10-30% of the expression of NbAIII.15 (**Figure 4-4C and D**). However, when we compared the ratio of expression of these nanobodies variants in the UCH37 KO cells to the WT, we discovered that the protein level of the nanobodies are higher in the UCH37 KO cells (**Figure 4-4E**). This result contradicts our hypothesis, as we expected the quantities of the nanobodies containing destabilizing mutations are lower in the absence of UCH37. We reasoned that since UCH37 plays a crucial role in regulating the proteasomal degradation, the protein turnover of the nanobodies expressed in the UCH37 KO cells could be suppressed and comparison of the nanobodies level between the WT and UCH37 KO cells could be misleading.

Here, we decided to perform a rescue experiment. We either co-transfected the UCH37 KO cells with NbAIII.15 variants and UCH37 or transfected the UCH37 KO cells with just NbAIII.15 variants. The expression level of NbAIII.15 variants under both conditions were then compared (**Figure 4-5A and Figure B-13**). We observed that for most variants showed an approximately 5 folds of induction in the presence of UCH37 (**Figure 4-5B**). Next, we questioned whether the introduced mutations would affect the interactions of NbAIII.15 with proteasome associated UCH37. Immunoprecipitations using transiently expressed conditionally stable NbAIII.15 were performed. Same as NbAIII.15, these variants are capable to immunoprecipitated with proteasome associated UCH37 and other proteasomal subunits (**Figure 4-6**). We decided to focus on NbAIII.15 S70R/S121F as our conditionally stable nanobody for our

downstream experiments. In addition, we performed the similar rescue experiment using NbCtrl and NbCtrl S70R/S121F. Western blot showed that the fold of induction of the NbCtrl S70R/S121F is around 2-fold, suggesting that it was not sensitive to the expression of UCH37 (**Figure B-14**). This further validates that the NbAIII.15 S70R/S121F is conditionally stabilized by UCH37.

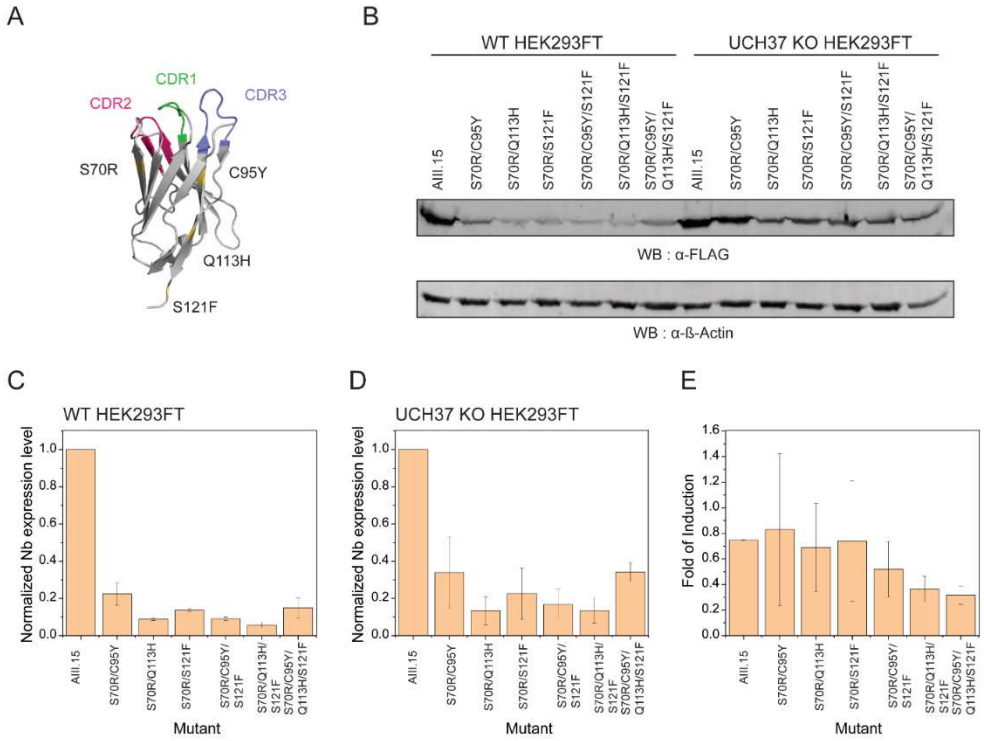

Figure 4-4: Conditionally stable NbAIII.15 in WT and UCH37 KO HEK293FT cells

(**A**) Crystal structure of nanobody (PDB ID: 5VNV) with the reported destabilizing mutations at S70R, C95Y, Q113H, S121F. The CDRs are highlighted in color (Green for CDR1, pink for CDR2, and blue for CDR3). (**B**) Representative Western blot showing the protein level of NbAIII.15 variants in the WT and UCH37 KO HEK293FT cell lysates. Expression level of NbAIII.15 variants containing destabilized mutations normalized to NbAIII.15 in (**C**) WT and (**D**) UCH37 KO HEK293FT cells. Error bars represent the standard error of two biological replicates. (**E**) Fold of induction of NbAIII.15 variants in WT HEK293FT cells compare to UCH37 KO HEK293FT cells. Error bars represent the standard error of two biological replicates.

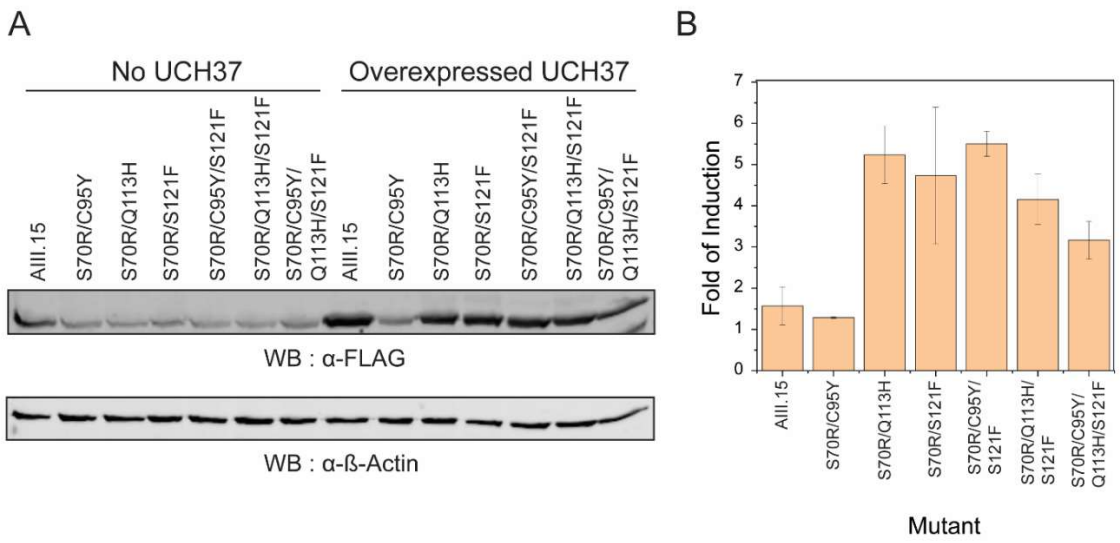

Figure 4-5: Conditionally stable NbAIII.15 in UCH37 KO HEK293FT cells without and with ectopically expressed UCH37

(**A**) Representative Western blot of the cell lysates and (**B**) fold of induction of NbAIII.15 variants in UCH37 KO HEK293FT cells without and with overexpressed UCH37. Error bars represent the standard error of two biological replicates.

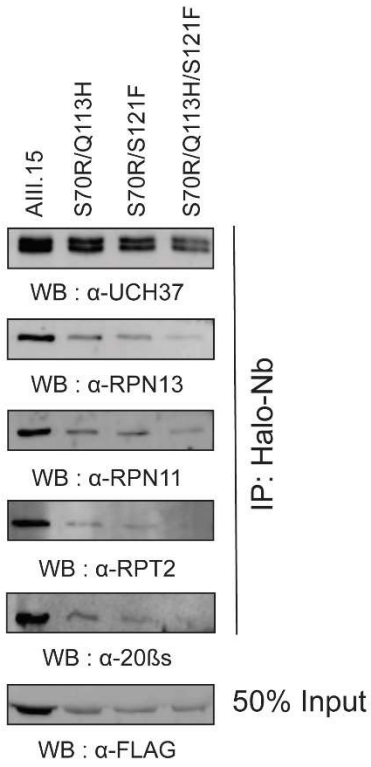

Figure 4-6: Immunoprecipitation of endogenous UCH37 by intracellularly expressed conditionally stably nanobodies.

Representative Western blots showing the transient expression of nanobodies and the immunoprecipitation of UCH37, RPN13, RPN11, and RPT2, which are the 19S RP subunits and 20βs, which is the 20S CP subunits by intracellularly expressed NbAIII.15 variants.

Discussion and Future Directions

In Chapter 3, we have identified nanobodies specific towards proteasome associated UCH37. In this study, we performed a detailed biochemical characterization of the nanobodies. UCH37 is reported to preferentially hydrolyze the Lys48 linked branched ubiquitin chains; we showed that the nanobodies display varied inhibitory effect towards the debranching activity of the UCH37 in different complex. Although most nanobodies inhibit the UCH37·RPN13 with an IC_{50} at low micromolar range, they could not inhibit the proteasome-associated UCH37. We have established that the nanobodies bind at the exosite of UCH37·RPN13$_{DEU}$. One of the possible reasons is the recruitment of UCH37·RPN13 to proteasome modulates or sterically blocks the epitope recognized by the nanobodies. Detailed structural analysis is needed to explain the observed discrepancy in the inhibitory effect of UCH37·RPN13 and proteasomal UCH37.

As the nanobodies do not affect the debranching activities of proteasomal UCH37 *in vitro*, they could be developed as an imaging tool to probe the proteasome-bound UCH37 with minimal perturbation towards its catalytic activity. The major technique in visualizing the UCH37 is through immunofluorescence, which requires cell fixation and is prone to the introduction of artefacts. While another common tagging practices is the generation of UCH37 fused to fluorescent protein, our group has shown that the N and C terminal of UCH37 are important in the activity of UCH37.[29,30] Here, using nanobodies targeting UCH37 could circumvent these issues; it also permits live cell imaging, which is crucial to

study the dynamic recruitment of UCH37. We do recognize one of the potential pitfalls is the overexpression of the ectopically expressed nanobodies, which may complicate the detection of the localization of UCH37 on the proteasome. We have further engineered our most potent proteasome-associated UCH37 nanobody, NbAIII.15 to a conditionally stable nanobody. [24] Here, the conditionally stable NbAIII.15 (csAIII.15) is constantly destabilized unless it is recognized by UCH37. csAIII.15 has low expression *in cellulo*, which might facilitate the detection of proteasomal UCH37 at endogenous level. We also verified that csAIII.15 interacted with proteasome associated UCH37.

Generation of cell line stably expressing Halo-csAIII.15 under inducible promoter is now underway to further control the expression of the imaging probe.[31] Here, we proposed to label the induced csNbAIII.15 with the cell-permeable HaloTag TAMRA ligand. Simultaneously, we would label the proteasome by treating the cells with a proteasome activity-based probe, Me_4Bodipy-Ahx$_3$-L$_3$-VS, which targets the β subunits of the 20S CP (**Figure 4-7**).[32] Colocalization analysis would be performed to investigate the recruitment of UCH37 on proteasome.

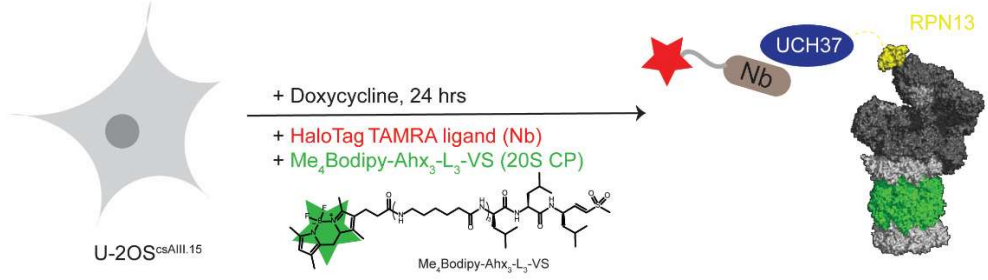

Figure 4-7: Strategy to visualize proteasomal UCH37 and 20S CP

Materials and Methods

Plasmid and cloning

His_8-MBP-TEV-UCH37 *(Isoform 3)* in pVP16 vector was obtained from DNASU. His_{10}-RPN13 in pET19b was a gift from Joan Conaway & Ronald Conaway (Addgene plasmid # 19423). All PCR amplifications were performed with Phusion® High-Fidelity DNA polymerase (New England Biolabs). The primers used in this study were purchased from IDT DNA. All the sequences were verified through Sanger sequencing. The cloning of His_6-nanobodies and HA-FLAG-HaloTag-nanobodies were described in the Materials and Methods in Chapter 3. The point mutations were introduced into the nanobody through site-directed mutagenesis using megaprimer PCR.

Cell culture

HEK293 stably expressing RPN11-His_6-TEV-biotin tag-His_6 ($HEK293^{RPN11\text{-}HTBH}$), HEK293FT and UCH37 KO HEK293FT were cultured in DMEM medium (Genesee Scientific) supplemented with 10% fetal calf serum (Genesee Scientific), 100 units/ml penicillin and 100 µg/ml streptomycin at 37°C in a humified atmosphere with 5% CO_2.

Protein purification

All protein purifications were performed at 4°C. For the purification of UCH37, UCH37·RPN13 and His_6-nanobodies, refer to the Materials and Methods in Chapter 3.

To prepare K6/K48 linked high molecular weight (HMW) ubiquitin chains, 0.5 mM Ub were reacted with 0.2 μM E1, 2 μM UbcH7, 1 μM NleL in a buffered solution containing 40 mM Tris, pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM ATP, 6 mM DTT at 37°C overnight. To prepare biotinylated K6/K48 linked HMW chains, 2 mM Ub were reacted with 0.2 μM E1, 2 μM UbcH7, 1 μM NleL in a buffered solution containing 40 mM Tris, pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM ATP, 6 mM DTT at 37°C for 3 hrs before spiking in 20 μM of biotinylated Ub (The preparation of the biotinylated Ub was described in the Materials and Methods in Chapter 2). The chain synthesis reactions was further continued for overnight. The ubiquitination reactions were quenched with 10 mM DTT in 50 mM NH_4OAc, pH 4.4 solution for at least 30 mins. The HMW chains was further purified through cationic exchange chromatography using MonoS on an ÄKTA purifier (GE Healthcare). The equilibrate buffer was 50 mM NH_4OAc, pH 4.4, 100 mM NaCl and the elution buffer was 50 mM NH_4OAc, pH 4.4, 1 M NaCl. For the biotinylated Ub chain, Western blot was performed using streptavidin, Alexa Fluor™ 647 conjugate (ThermoFisher Scientific) at 1:15000 dilution to confirm the generation of biotinylated HMW chains.

USP14 depleted 26*S* proteasome was purified through affinity purification from HEK293 cells stably expressing RPN11-His_6-TEV-biotin tag-His_6 ($HEK293^{RPN11\text{-}HTBH}$). Twenty 15 cm^2 dishes of $HEK293^{RPN11\text{-}HTBH}$ cells were cultured, harvested and lysed in 50 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 1 mM DTT, 10% glycerol. The HTBH-tagged RPN11 containing proteasome was immobilized for overnight at 4°C on streptavidin agarose resin (Genscript). The

immobilized proteasome was washed extensively to remove USP14 with 40 mM HEPES, pH 7.5, 200 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 1 mM DTT, 10% glycerol before eluted by treating the resin with 2 μM TEV enzyme for 1.5 hrs at room temperature. The eluted proteasome was stored in 40 mM HEPES, pH 7.5, 200 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 10% glycerol. Western blot was performed to investigate the composition of the purified proteasome using the following antibodies: rabbit anti-UCH37 [EPR4896] (abcam), rabbit anti-ADRM1 (D9Z1U) (Cell Signaling Technology), rabbit anti-USP14 (D8Q6S) (Cell Signaling Technology), rabbit anti-PSMD14 [EPR4258] (abcam) and mouse anti-PSMB7 (R&D systems) at 1:1000 dilution. IRDye® 680RD goat anti-rabbit and IRDye® 800CW goat anti-mouse (LI-COR Biosciences) were used as the secondary antibodies at 1:15000 dilution. Especially, western blot against rabbit anti USP14 was to verify the absence of USP14 from the purified proteasome. The activities of the 20 *S* proteasome subunit and the proteasome associated deubiquitinase were verified using fluoregenic proteasome substrate, suc-LLVY-AMC (Boston Biochem) and fluoregenic ubiquitin hydrolase substrate, Ub-AMC (Boston Biochem).

Deubiquitinase assay

For the UCH37 and UCH37·RPN13 cleavage assay, 1 μM of the enzyme was pre-incubated with nanobody at a concentration gradient in a buffered solution of 50 mM HEPES, pH 7.5, 50 mM NaCl, 3 mM GSH before reacted with 0.3 mg/ml HWM K6/K48- linked ubiquitin chains for 30 mins at 37^{o}C. The activity of the

UCH37 and UCH37·RPN13 in the presence of nanobody was accessed by gel densitometry. Gels were stained with SYPRO™ Ruby and scanned on a Typhoon FLA 9500 (GE Healthcare). Data analysis was performed on Image Studio Lite (LI-COR Biosciences). The IC_{50} was determined from the direct fitting to a dose response model on Origin.

For the proteasome-associated UCH37 cleavage assay, USP14 depleted proteasome (final concentration was 40 nM) were incubated with serial diluted Nb for 1 hr at 37°C before reacted with 0.3 mg/ ml of biotinylated K6/K48 HWM Ub chains for 3 hrs at 37°C in a buffered solution containing 50 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 3 mM GSH. The reactions were quenched by SDS loading dye prior to western blot analysis using streptavidin, Alexa Fluor™ 647 conjugate (ThermoFisher Scientific) at 1:15000 dilution. Data acquisition and analysis was performed on Image Studio Lite (LI-COR Biosciences).

Transient transfection

Transient transfections of mammalian cells were performed using Xfect™ transfection reagent (Takara Bio) according to the manufacturer's protocol and scaling recommendation. Briefly, HEK293FT were grown to 35% confluency. Plasmid DNA was incubated with the Xfect polymer at room temperature for at least 10 mins before added drop wise to the adhered cells. The cells were incubated with the complete growth media containing DNA nanoparticles for 4

hrs before replacing it with fresh media and further incubated for 48 hrs and the cells were prepared for downstream analysis.

Analysis for conditionally stable nanobodies

12 well plates of HEK293FT and UCH37 KO HEK293FT cells were grown and transfected with respective conditionally stable Nbs as described above. The cells were harvested and lysed in RIPA buffer. Western blot analysis was performed using chicken anti-FLAG (Exalpha Biologicals) at 1:5000 dilution and rabbit anti-β-actin (ab8827) (Abcam) at 1:3000 dilution. Goat anti chicken IgY (H+L) secondary antibody, Alexa Fluor™ 647 (ThermoFisher Scientific) and IRDye® 800CW goat anti-rabbit (LI-COR Biosciences) were used as the secondary antibodies at 1:15000 dilution. Data acquisition and analysis was performed on Image Studio Lite (LI-COR Biosciences).

References

1. Voges, D., Zwickl, P. & Baumeister, W. The 26S Proteasome: A molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* **68**, 1015–1068 (1999).
2. Tanaka, K. The proteasome: Overview of structure and functions. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* **85**, 12–36 (2009).
3. Livneh, I., Cohen-Kaplan, V., Cohen-Rosenzweig, C., Avni, N. & Ciechanover, A. The life cycle of the 26S proteasome: From birth, through regulation and function, and onto its death. *Cell Res.* **26**, 869–885 (2016).
4. Collins, G. A. & Goldberg, A. L. The Logic of the 26S Proteasome. *Cell* **169**, 792–806 (2017).
5. Martinez-Fonts, K. *et al.* The proteasome 19S cap and its ubiquitin receptors provide a versatile recognition platform for substrates. *Nat. Commun.* **11**, (2020).

6. Hamazaki, J., Hirayama, S. & Murata, S. Redundant Roles of Rpn10 and Rpn13 in Recognition of Ubiquitinated Proteins and Cellular Homeostasis. *PLoS Genet.* **11**, 1–20 (2015).

7. de Poot, S. A. H., Tian, G. & Finley, D. Meddling with Fate: The Proteasomal Deubiquitinating Enzymes. *J. Mol. Biol.* **429**, 3525–3545 (2017).

8. Lee, B.-H. *et al.* USP14 deubiquitinates proteasome-bound substrates that are ubiquitinated at multiple sites. *Nature* **532**, 398–401 (2016).

9. Deol, K. K., Crowe, S. O., Du, J., Bisbee, H. & Strieter, E. R. Proteasome-Bound UCH37 Debranches Ubiquitin Chains to Promote Degradation 1 2. doi:10.1101/2020.02.21.960088.

10. Fabre, B. *et al.* Label-free quantitative proteomics reveals the dynamics of proteasome complexes composition and stoichiometry in a wide range of human cell lines. *J. Proteome Res.* **13**, 3027–3037 (2014).

11. Wang, X. & Huang, L. Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. *Mol. Cell. Proteomics* **7**, 46–57 (2008).

12. Koulich, E., Li, X. & N. DeMartino, G. Relative Structural and Functional Roles of Multiple Deubiquitylating Proteins Associated with Mammalian 26S Proteasome. *Mol. Biol. Cell* **19**, 1072–1082 (2008).

13. Wang, X., Yen, J., Kaiser, P. & Huang, L. Regulation of the 26S proteasome complex during oxidative stress. *Sci. Signal.* **3**, ra88–ra88 (2010).

14. Motosugi, R. & Murata, S. Dynamic regulation of proteasome expression. *Front. Mol. Biosci.* **6**, 4–11 (2019).

15. Berko, D. *et al.* Inherent asymmetry in the 26S proteasome is defined by the ubiquitin receptor RPN13. *J. Biol. Chem.* **289**, 5609–5618 (2014).

16. Lundgren, J., Masson, P., Realini, C. A. & Young, P. Use of RNA Interference and Complementation To Study the Function of the Drosophila and Human 26S Proteasome Subunit S13. *Mol. Cell. Biol.* **23**, 5320–5330 (2003).

17. Chadchankar, J. *et al.* Inactive USP14 and inactive UCHL5 cause accumulation of distinct ubiquitinated proteins in mammalian cells. *PLoS One* **14**, 1–23 (2019).

18. Traenkle, B. & Rothbauer, U. Under the microscope: Single-domain antibodies for live-cell imaging and super-resolution microscopy. *Front. Immunol.* **8**, 1–8 (2017).

19. Keller, B. M. *et al.* Chromobodies to quantify changes of endogenous protein concentration in living cells. *Mol. Cell. Proteomics* **17**, 2518–2533 (2018).

20. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–898 (2018).

21. Tachie-Menson, T. *et al.* Characterisation of the biochemical and cellular roles of native and pathogenic amelogenesis imperfecta mutants of FAM83H. *Cell. Signal.* **72**, 109632 (2020).

22. D'Arcy, P. *et al.* Inhibition of proteasome deubiquitinating activity as a new cancer therapy. *Nat. Med.* **17**, 1636–1640 (2011).

23. Gupta, A. *et al.* Facile target validation in an animal model with intracellularly expressed monobodies. *Nat. Chem. Biol.* 1–6 (2018) doi:10.1038/s41589-018-0099-z.

24. Tang, J. C. Y. *et al.* Detection and manipulation of live antigen-expressing cells using conditionally stable nanobodies. *Elife* **5**, 1–27 (2016).

25. Sahtoe, D. D. *et al.* Mechanism of UCH-L5 Activation and Inhibition by DEUBAD Domains in RPN13 and INO80G. *Mol. Cell* **57**, 887–900 (2015).

26. VanderLinden, R. T. *et al.* Structural Basis for the Activation and Inhibition of the UCH37 Deubiquitylase. *Mol. Cell* **57**, 901–911 (2015).

27. Al-Lazikani, B., Lesk, A. M. & Chothia, C. Standard conformations for the canonical structures of immunoglobulins. *J. Mol. Biol.* **273**, 927–948 (1997).

28. Ariotti, N. *et al.* Ultrastructural localisation of protein interactions using conditionally stable nanobodies. *PLoS Biol.* **16**, 1–11 (2018).

29. Los, G. V. *et al.* HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008).

30. Meerbrey, K. L. *et al.* The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3665–3670 (2011).

31. Leestemaker, Y. *et al.* Proteasome Activation by Small Molecules. *Cell Chem. Biol.* **24**, 725-736.e7 (2017).

32. Schipper-Krom, S. *et al.* Visualizing Proteasome Activity and Intracellular Localization Using Fluorescent Proteins and Activity-Based Probes. *Front. Mol. Biosci.* **6**, 1–18 (2019).

33. Buchfellner, A. *et al.* A new nanobody-based biosensor to study endogenous PARP1 in vitro and in live human cells. *PLoS One* **11**, 1–23 (2016).

34. Stadler, C. *et al.* Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nat. Methods* **10**, 315–323 (2013).

35. Albert, S. *et al.* Direct visualization of degradation microcompartments at the ER membrane. *Proc. Natl. Acad. Sci.* 1–12 (2019) doi:10.1073/pnas.1905641117.

Appendix A

GENERATION OF A SYNTHETIC NANOBODY LIBRARY

Nanobody is a single domain antibody, with the size around 15 kDa. Similar to the conventional antibody, nanobody recognizes its epitope using its complementary determining regions. In contrast to antibody, nanobody is 10 times smaller and readily assembled to functional module *in cellulo*; hence it has been widely applied to study numerous protein intracellularly.[1] The engineer of nanobody could be achieved through either immunization of the camelids, such as llama or alpaca with the antigen to elicit immune response. The DNA sequence of the nanobodies were then identified through reverse transcription of the extracted mRNA from the lymphocytes. The immunized libraries were then subjected to biopanning, mainly through phage display. On the other hand, several groups had reported the generation of synthetic nanobody libraries and successfully isolated nanobody *in vitro* through directed evolution.[2,3] These *in vitro* platform is ideal to isolate conformational selective nanobodies and if the antigen is sensitive to the environment. In addition, the synthetic nanobodies libraries could be cloned into different display platform, such as yeast display. The major advantage of yeast display is it could be coupled to fluorescent activated cell sorting (FACS). The multicolor flow cytometry enables the normalization of the antigen binding signals with the level of displayed nanobody and with FACS, each clone could be scrutinized in a high throughput screen.

In this chapter, we demonstrated the preparation and characterization of a yeast display-based synthetic nanobody library based on an established library

design.[2] We have used the library to isolate nanobodies targeting two deubiquitinases from distinct families: USP30 and UCH37.

USP30 is a deubiquitinase belongs to the USP family. Studies have elucidated that USP30 anchors at the outer membrane mitochondria and antagonizes mitophagy.[4] The inhibition of USP30 promotes the clearance of damaged mitochondria and maybe beneficial for Parkinson's disease.[5] It is further demonstrated that mitochondrial damage triggers the ubiquitination of an atypical Lys6- linked ubiquitin chain on the mitochondrial substrate, which is hydrolyzed by USP30.[6] Structural study of USP30 showed that it has the preference towards the Lys6-linked ubiquitin chains is through a unique ubiquitin binding surface.[7,8,9] We reasoned that a USP30 nanobody could be used to probe USP30 spatially and temporally to give insights into the physiological roles of USP30 in neurological disease. (The generation and characterization of UCH37 nanobodies are described in Chapter 3 and 4).

Generation of a synthetic Nb library

By adopting the established library design, we displayed our synthetic nanobody library on the yeast surface with the yeast Aga1p-Aga2p display system.[2] Our ensemble of nanobodies consist similar thorough randomization at the CDRs and partial randomization at the residues adjacent to the CDRs. The randomization sites are: (i) residues 27-33, (ii) residues 50-58 and (iii) residues 94-100 according to the Chothia numbering.[10,11] Here, the partial randomizations are achieved through degenerate codons, while the full randomizations are attained by the utilization of NNK codons, which code for all twenty amino acids

and one stop codon. We acknowledged that the usage of NNK codons will introduce Cys with undesired chemical property into our nanobody library. At the same time, there is a 5% risk to come across a stop codon in the sequence which lead to the generation of a non-functional truncated nanobody. Nonetheless, we proceeded with the NNK codons from an economic standpoint.

Nanobody is characterized with its long CDR3 with variable length.[12,13,14] Studies have shown that the diversities of the CDR3 are crucial in distinguishing the target antigens, and their different lengths determine the binding modes and the theoretical sequence space of the nanobodies.[15,3,16] Here, similar to the established nanobody library, we introduced 3 length variations at the CDR3, with 8, 10, and 13 amino acid residues (according to the Chontia numbering). Studies have shown that CDR3 with short length, such as 8 and 10 residues mainly adopts a concave interaction surface, and CDR3 with medium length, for example, 13 amino acids will offer an extensive loop interaction surface. While longer CDR3, with more than 17 residues and binds with convex surface, is not considered in our library design, as the probability of encountering stop codon would be too high.

The nanobody library is incorporated into the yeast surface display vector with N terminus Aga2p protein and is flanked between a HA tag and c-Myc tag (**Figure A-1A**). Upon transformation, we obtained around 7 x 10^7 transformants. The utilization of NNK codons would result in 40-50% of truncated nanobodies; we have performed flow cytometry analysis of the induced library, using the anti-HA

and anti-cMyc and confirmed the percentage of the full length nanobodies were 44% of the expressed nanobodies (**Figure A-1B**).

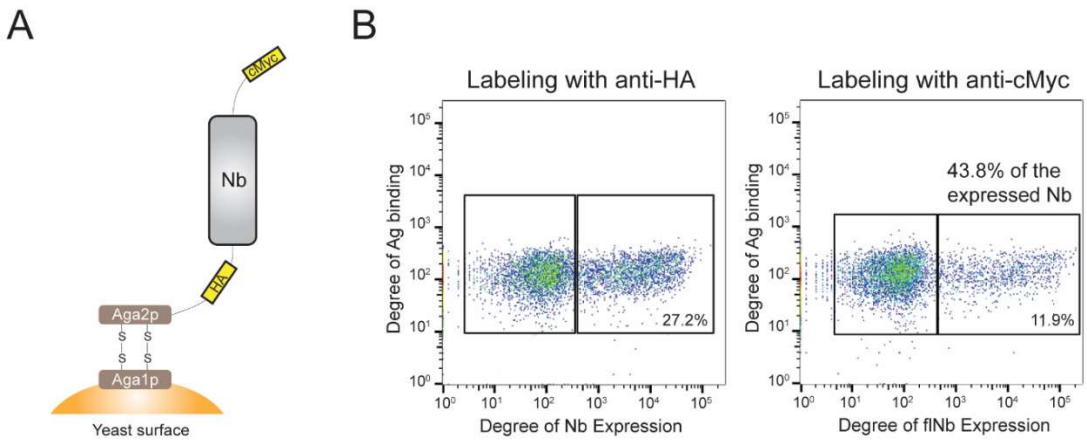

Figure A-1: Yeast display of the synthetic nanobody library.

(**A**) Scheme of the displayed nanobody library on the yeast surface. (**B**) Flow cytometry dot plots of the induced nanobody library. The displayed nanobodies are either detected with the labeling at the HA tag fused to the N terminus of nanobody to identify the degree of induction (Left) or detected with the labeling at the cMyc tag fused to the C terminus of nanobody to determine the degree of the expression of full length nanobody (Right).

We further validated the quality of the library using next generation sequencing (NGS). Here, we used the Illumina's MiSeq paired-end sequencing (2 x 250 bps) which covers the entire nanobody. We recognized that the capacity of our MiSeq run is 1 million read at maximum, which is below the size of our library and the characterization of each clone could not be achieved. Nonetheless, the sampling depth using NGS is significantly higher than the traditional Sanger sequencing and could offer a more accurate assessment of the library diversity.[17] From our MiSeq sequencing, among 608168 reads, 33% of

the clones were in frame. 73% of the in-frame ensembles did not contain any stop codon with 93% of the full length nanobodies were distinct (clones with unique sequence) (**Table A-1**). Thereby, the diversity of our nanobody library is estimated to be around 2 x 10^7 distinct clones. A closer inspection of the full length nanobodies showed that the ratio of the nanobodies with 8, 10, and 13 CDR3 amino acid residues are 3:2:1 (**Figure A-2**). In addition, we interrogated the diversities of the amino acids at the CDR loops and our observed frequencies resembled the theoretical frequencies, suggesting the unbiased incorporation of the degenerate codons (**Figure A-3**)

.

Table A-1: Overview of sequencing data of nanobody library

DESCRIPTION	NO. OF READS	PERCENTAGE (%)	RELATIVE PERCENTAGE (%)
RAW	608168	100.0	100.0
IN FRAME	201453	33.1	33.1
FULL LENGTH	146231	24.0	72.6
UNIQUE	136349	22.4	93.2

Raw data represents the total number of merged reads with low quality data removed; in frame is the number of reads containing in frame nanobody framework; the number of reads of full length is determined after removing the nanobody containing at least one stop codon; and unique read is the number of nanobody with unique sequence. The percentage is calculated based on the number of reads compare to the raw data while the relative percentage is based on the percentage relative to previous sub-sampled reads.

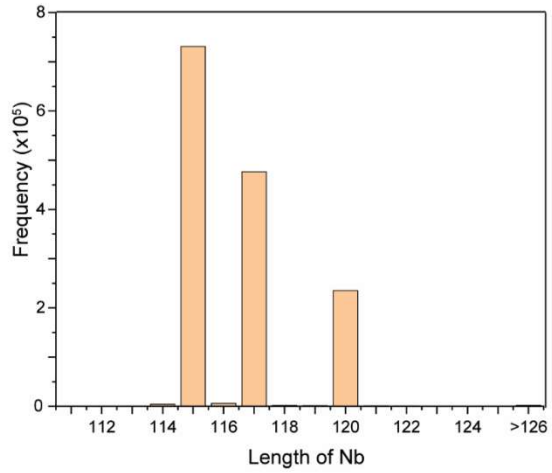

Figure A-2: NGS Analysis of the length of nanobodies library

Barplot showing the distribution of the ensemble of nanobodies based on their length.

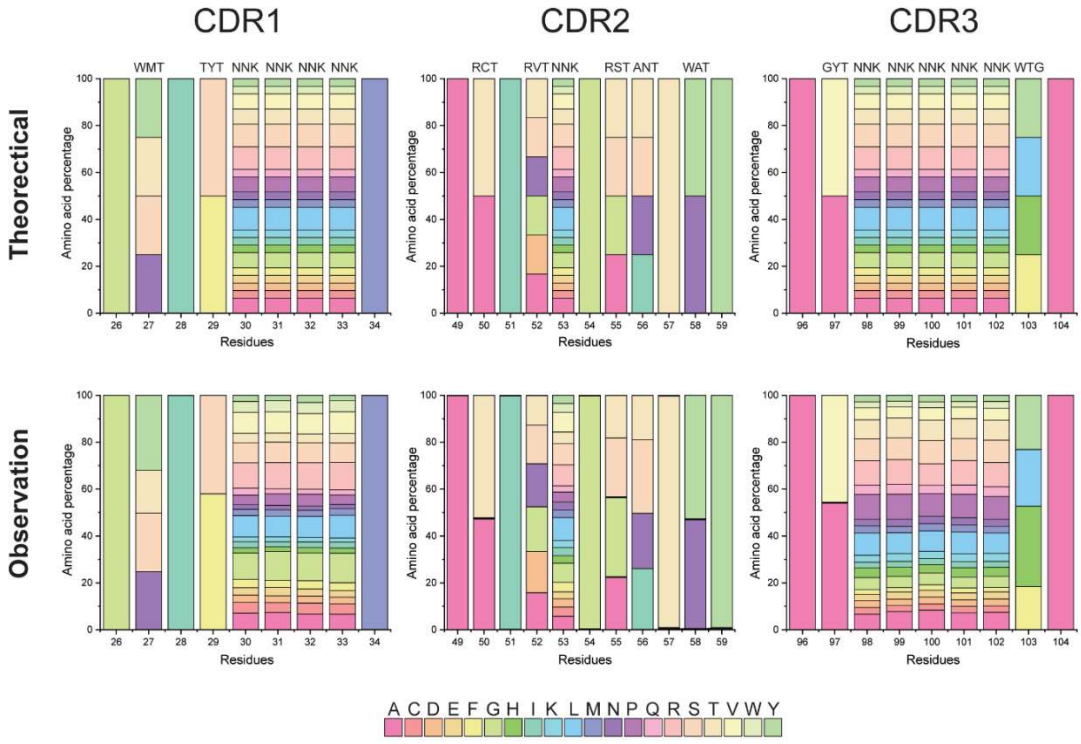

Figure A-3: Side by side comparison of the theoretical and observed amino acid percentage of the nanobody library focusing at CDRs.

The amino acid percentage of the CDR1 (Left) and CDR2 (Middle) of the full length nanobodies were calculated. As the nanobodies library consist of members with different CDR3 length, only the ensembles with 8 amino acid residues at CDR3 (Right) was used in this analysis.

Generation of the USP30 Nbs

We reasoned that a nanobody which probe intracellular USP30 would provide information of the functional role of USP30 in a cellular context. Our preliminary screening against ubiquitin conjugated to USP30 had enriched three different low affinity nanobody clones towards USP30, as suggested from the flow cytometry assay (**Figure A-4A and B**). Especially one nanobody, NbBI.1 showed inhibitory effect on the Lys6 linked- ubiquitin chain hydrolysis activity of USP30 (**Figure A-4C**). Affinity maturation is currently in the progress to improve the binding parameters of these nanobodies.

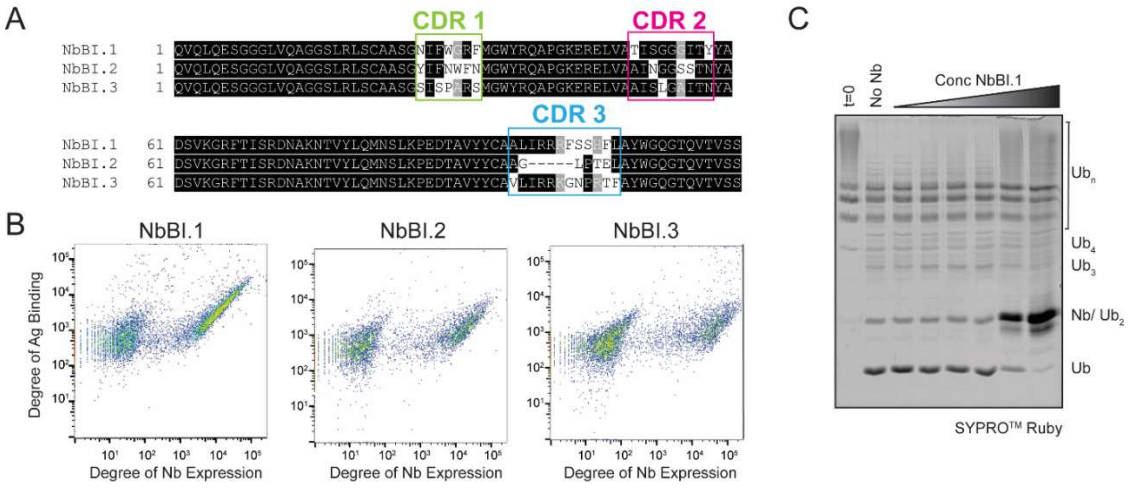

Figure A-4: Isolating nanobodies for USP30.

(**A**) Amino acid sequences of the enriched nanobodies for USP30. The diversified residues in the CDRs are highlighted in colored box. (**B**) Flow cytometry plots showing bindings of yeast displayed- nanobodies towards 316 nM USP30. (**C**) Lys6 linked-ubiquitin chain cleavage assay by USP30 in the presence of NbBI.1 in a concentration gradient.

The capability to isolate of nanobodies for different antigens suggested that this nanobody library is robust. We recognized that the main limitation of our

library is the existence of stop codon due to the usage of NNK codon as a method to introduce full randomization. This had inevitably led to the generation of truncated nanobody; it also prevented us from preparing nanobody ensembles with longer CDR3 to reduce the probability of encountering stop codon in that region. Usage of the oligos with trimer phosphoramidites would circumvent this issue.[18]

Materials and Methods

Cell culture

For yeast cells used in the yeast surface display, referred to the cell culture section in the Material and Methods in Chapter 2.

Plasmid and Cloning

His_6-SUMO-$USP30_{67-517}$ in pOPINS vector was cloned from Flag-HA-USP30, which was a gift from Wade Harper (Addgene plasmid # 22578). All PCR amplifications were performed with Phusion® High-Fidelity DNA polymerase (New England Biolabs). The primers used in this study were purchased from IDT DNA. All the sequences were verified through Sanger sequencing.
For bacterial expression of the nanobody, the enriched nanobodies in the selections were cloned into pET28b (Novagen) through megaprimer PCR.

Construction of a synthetic Nanobody (Nb) library

Our Nb library design is based on published synthetic Nb library and is displayed through the conventional yeast surface display platform.[2,16,19] The Nb library was

constructed by overlapping extension PCR of five oligonucleotides which composed the framework structure (oligonucleotides a and e) and randomized regions (oligonucleotides b, c, and d_{short}/ d_{medium}/ d_{long}, and e) of Nb (**Table I-2**). The oligonucleotides were generated by performing PCR using a1_f+a2_r, b3_f+b4_r, c5_f+c6_6, d7_f+d8_short_r/ medium_r/ long_r, and e9_f+e10_r. The resulting mixture (oligonucleotides a, b, c, d, and e) were mixed at equal ratio for the following overlapping extension PCR; among them, oligonucleotides d were further composed of d_{short}, d_{medium}, and d_{long} mixed at a ratio of 1:1:2. Both PCR were performed using Phusion® high fidelity DNA polymerase (New England Biolabs) and the conditions were initial denaturation at 98^{o}C for 30 s, followed by 25 cycles of 98^{o}C for 10 s, 60^{o}C for 30 s, 72^{o}C for 30 s, with a final extension at 72^{o}C for 5 mins. The combined Nb gene library pool was further amplified using pCT_nb_PstI_f and pCT_nb_BamHI_r.(**Table I-2**) To generate Nb library, a yeast surface display vector containing nanobody, pCT_NbNO1 was first generated by inserting the amplified Nb gene into the plasmid at NheI and BamHI (New England Biolabs) restriction site. It was followed by transforming the yeast cells using 1 µg of the Bpu10I and BamHI (New England Biolabs) digested pCT_NbNO1 plasmid and 5 µg of the amplified Nb gene library pool through electroporation. The cells were recovered and the size of the transformants was estimated through the aforementioned procedure. The diversity of the Nb library was further validated through next generation sequencing (Refer to Material and Method section in Chapter 3 for NGS library sample preparation and analysis).

Table A-2: Primers used in the generation of the nanobody library

Primers	Sequence (5'-3')
a1_f	GCTAGCCAGGTGCAGCTGCAGGAAAGCGGCGGCGGCCTGGTGCAGGCGGGC
a2_r	GCCGCTCGCCGCGCAGCTCAGGCGCAGGCTGCCGCCCGCCTGCAC
b3_f	CGCGGCGAGCGGC**WMT**ATT**TYTNNKNNKNNKNNK**ATGGGCTGGTAT
b4_r	AAT**AGY**AGCAACAAGTTCGCGTTCTTTGCCCGGCGCCTGGCGATACCAGCCCAT
c5_f	GTTGCT**RCT**ATT**RVTNNK**GGT**RSTANT**ACC**WAT**TATGCGGATAGCGTGAAAGGCCGCTTT
c6_r	TTCATCTGCAGATACACGGTGTTTTTCGCGTTATCGCGGCTAATGGTAAAGCGGCCTTT
d7_f	TCTGCAGATGAACAGCCTGAAACCGGAAGATACCGCGGTGTATTATTGCGCGG
d8_short_r	CCCCAATATGC**AWRMNNMNNMNNMNNMNNARC**CGCGCAATAAT
d8_medium_r	CCCCAATATGC**AWRMNNMNNMNNMNNMNNMNNMNNARC**CGCGCAATAAT
d8_long_r	CCCCAATATGC**AWRMNNMNNMNNMNNMNNMNNMNNMNNMNNMNNARC**CGCGCAATAAT
e9_f	GCATATTGGGGCCAGGGCACCCAGGTGACCGTGAGCAGCGGTGGTGGTGGTAGC
e10_r	GTCCCCGAAGTTCAGACCGGTGCTACCACCACCACCGCTGCTCACGGT
pCT_nb_PstI_f	GGTGGTGGTTCTGCTAGCCAGGTGCAGCTGCAGGAA
pCT_nb_BamHI_r	TTACAAGTCCTCTTCAGAAATAAGCTTTTGTTCGGATCCGCTACCACCACCGCT

Codons used for partial and full randomization are highlighted in bold.

Protein purification

All protein purifications were performed at 4°C.

Avi-USP30$_{67-517}$/USP30$_{67-517}$ were purified through nickel (Gold Biotechnology) affinity chromatography. In brief, the USP30 construct were expressed in *E. coli* Rosetta (DE3) pLysS. The cells were grown at 37°C to OD 0.5 before induced with 0.1 mM IPTG at 20°C overnight. The cells were lysed in 50 mM Tris, pH 8, 300 mM NaCl, 10 mM imidazole and subjected to nickel resin to pull down the

His_8-SUMO-USP30 construct. The protein was eluted with 50 mM Tris, pH 8, 150 mM NaCl, 300 mM imidazole, followed by ULP1 cleavage reaction in 50 mM Tris, pH 8, 150 mM NaCl, 1 mM TCEP to remove the His_8-SUMO tag. The protein was further purified through size exclusion chromatography on a HiLoad Superdex 75 pg on an ÄKTA purifier (GE Healthcare). The running buffer was 50 mM Tris, pH 7.5, 300 mM NaCl, 1 mM TCEP. For the preparation of the biotinylated Ub conjugated to $USP30_{67-517}$, biotinylated $USP30_{67-517}$, and His_6-nanobody, refer to the Material and Method section in Chapter 3.

To prepare K6-linked ubiquitin chain, 0.5 mM Ub K48R (referred to protein expression and purification section in the Materials and Methods in Chapter 2 for ubiquitin K48R purification) was reacted with 0.2 μM E1, 2 μM UbcH5c, 0.5 μM MleL in a buffered solution containing 40 mM Tris, 10 mM $MgCl_2$, 10 mM ATP, and 0.2 mM DTT at pH 7.5 for 37^{o}C overnight. The reactions were quenched using 10 mM DTT in 50 mM NH_4OAc, pH4.4 solution for at least 30 min before purified using size exclusion chromatography using a HiLoad Superdex 75 pg on an ÄKTA purifier (GE Healthcare). The running buffer was 50 mM Tris, pH 7.5, 300 mM NaCl, 1 mM DTT.

Selection of Nanobodies (Nbs) against $USP30_{67-517}$

Briefly, the induced Nb library were subjected to two rounds of magnetic activated cell sorting (MACS) using biotinylated Ub conjugated to $USP30_{67-517}$ and followed by two rounds of fluorescent activated cell soritng (FACS) using either biotinylated Ub conjugated to $USP30_{67-517}$ or biotinylated $USP30_{67-517}$ as

antigen. Refer to Material and Method section for the preparation of the yeast cells for sorting.

Validation of the enriched mutations

Yeast cells were transformed with the plasmid encoded for the nanobodies with enriched mutations (referred to Yeast growth and induction in the Materials and Methods in Chapter 2). The yeast cells were cultured, induced, and washed as previously described. For validation, the cells were incubated with the biotinylated $USP30_{67-517}$ in PBS for 30 mins, at room temperature with gentle rocking. The cells were then similarly labeled with primary and secondary antibodies as mentioned in the *Selection of nanobodies against* $USP30_{67-517}$ before subjected to flow cytometry analysis on a BD LSR Fortessa X20.

USP30 activity assay

1 µM USP30 was pre-incubated with nanobody at a concentration gradient in a buffered solution of 50 mM HEPES, pH 7.5, 50 mM NaCl, 3 mM DTT before reacted with 0.3 mg/ml high molecular weight Lys6- linked ubiquitin chains for 30 mins at 37^{o}C. The activity of USP30 in the presence of nanobody was accessed by gel densitometry. Gels were stained with SYPRO^{TM} Ruby and scanned on a Typhoon FLA 9500 (GE Healthcare). Data analysis was performed on Image Studio Lite (LI-COR Biosciences).

References

1. Dmitriev, O. Y., Lutsenko, S. & Muyldermans, S. Nanobodies as probes for protein dynamics in vitro and in cells. *J. Biol. Chem.* **291**, 3767–3775 (2016).
2. Mcmahon, C. *et al.* Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, (2018).
3. Zimmermann, I. *et al.* Synthetic single domain antibodies for the conformational trapping of membrane proteins. *Elife* **3**, 1–32 (2018).
4. Nakamura, N. & Hirose, S. Regulation of Mitochondrial Morphology by USP30, a Deubiquitinating Enzyme Present in the Mitochondrial Outer Membrane. *Mol. Biol. Cell* **19**, 1903–1911 (2008).
5. Bingol, B. *et al.* The mitochondrial deubiquitinase USP30 opposes parkin-mediated mitophagy. *Nature* **510**, 370–375 (2014).
6. Cunningham, C. N. *et al.* USP30 and parkin homeostatically regulate atypical ubiquitin chains on mitochondria. *Nat. Cell Biol.* **17**, 160–169 (2015).
7. Hospenthal, M. K., Freund, S. M. V. & Komander, D. Assembly, analysis and architecture of atypical ubiquitin chains. *Nat. Struct. Mol. Biol.* **20**, 555–565 (2013).
8. Gersch, M. *et al.* Mechanism and regulation of the Lys6-selective deubiquitinase USP30. *Nat. Struct. Mol. Biol.* **24**, 920–930 (2017).
9. Sato, Y. *et al.* Structural basis for specific cleavage of Lys6-linked polyubiquitin chains by USP30. *Nat. Struct. Mol. Biol.* **24**, 911–919 (2017).
10. Al-Lazikani, B., Lesk, A. M. & Chothia, C. Standard conformations for the canonical structures of immunoglobulins. *J. Mol. Biol.* **273**, 927–948 (1997).
11. Dunbar, J. *et al.* SAbPred: a structure-based antibody prediction server. *Nucleic Acids Res.* **44**, W474–W478 (2016).
12. Sircar, A., Sanni, K. A., Shi, J. & Gray, J. J. Analysis and Modeling of the Variable Region of Camelid Single-Domain Antibodies. *J. Immunol.* **186**, 6357–6367 (2011).
13. Mitchell, L. S. & Colwell, L. J. Comparative analysis of nanobody sequence and structure data. *Proteins Struct. Funct. Bioinforma.* **86**, 697–706 (2018).
14. Li, X. *et al.* Comparative analysis of immune repertoires between bactrian Camel's conventional and heavy-chain antibodies. *PLoS One* **11**, 1–15 (2016).

15. Xu, J. L. & Davis, M. M. Diversity in the CDR3 Region of V. *Immunity* **13**, 37–45 (2000).

16. Hackel, B. J., Kapila, A. & Wittrup, K. D. Picomolar Affinity Fibronectin Domains Engineered Utilizing Loop Length Diversity, Recursive Mutagenesis, and Loop Shuffling. (2008) doi:10.1016/j.jmb.2008.06.051.

17. Glanville, J. *et al.* Deep sequencing in library selection projects: What insight does it bring? *Curr. Opin. Struct. Biol.* **33**, 146–160 (2015).

18. Ono, A., Matsuda, A., Zhao, J. & Santi, D. V. The synthesis of blocked triplet-phosphoramidites and their use in mutagenesis. *Nucleic Acids Res.* **23**, 4677–4682 (1995).

19. Boder, E. T. & Wittrup, K. D. Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol.* **328**, 430–444 (2000).

Appendix B

SUPPLEMENTARY FIGURES AND TABLES

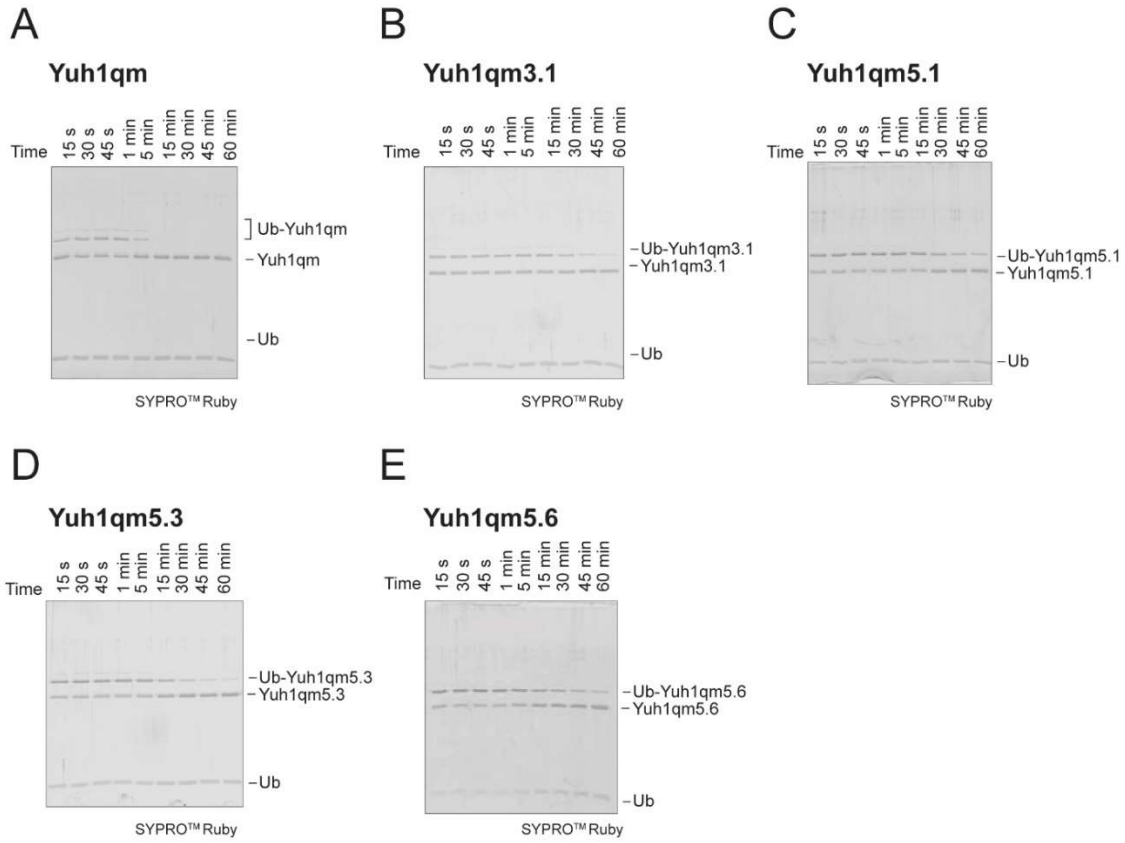

Figure B-1: Auto-ubiquitination of Yuh1 variants

Representative SDS PAGE analysis showing the whole gel of the autoubiquitination of (**A**) Yuh1qm, (**B**) Yuh1qm3.1, (**C**) Yuh1qm5.1, (**D**) Yuh1qm5.3 and (**E**) Yuh1qm5.6 with monoubiquitin.

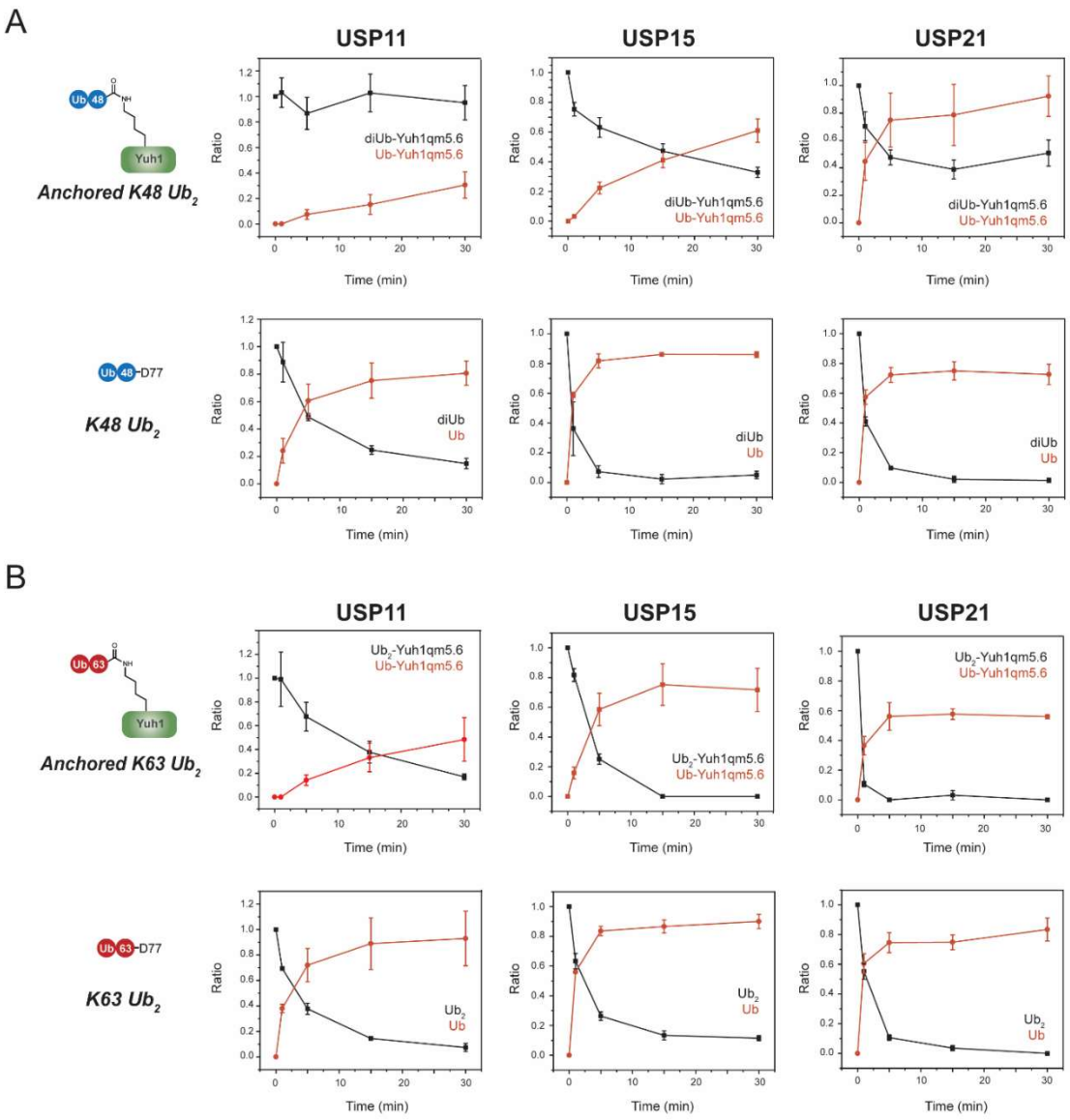

Figure B-2: Profiling USP activities towards anchored and free ubiquitin chains

Free and anchored (**A**) K48 Ub_2 and (**B**) K63 Ub_2 cleavage assay by USP11, USP15 and USP21 with plots showing the disappearance of anchored (or free) ubiquitin dimer and the formation anchored (or free) mono-ubiquitin. Error bars corresponded to standard error of three biological replicates.

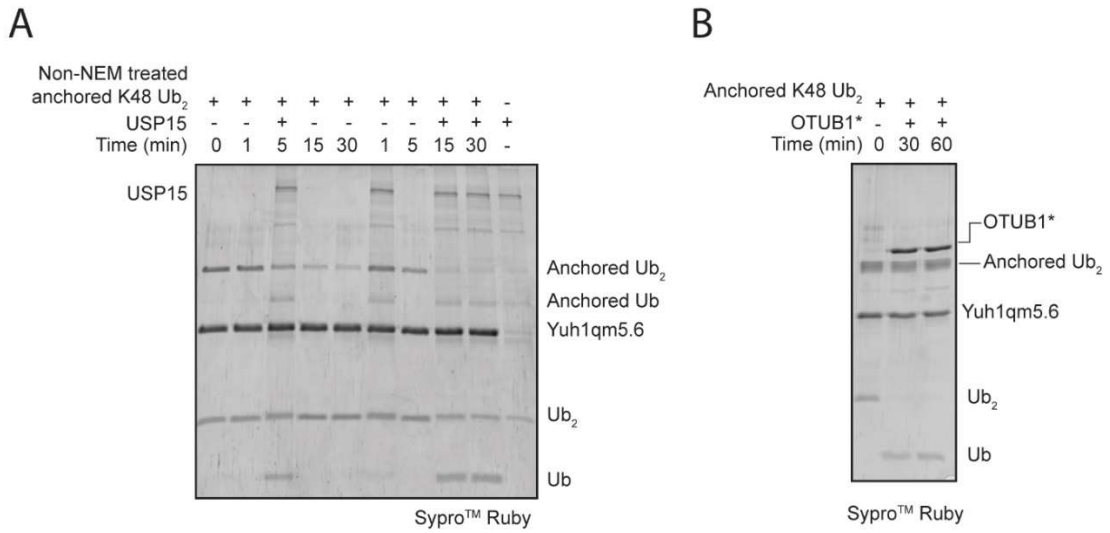

Figure B-3: Alternative anchored Lys48 linked ubiquitin dimer cleavage assay

(**A**) Non-NEM treated anchored Lys48 linked ubiquitin dimer cleavage assay by USP15. (**B**) Anchored and free Lys48 linked ubiquitin dimer cleavage assay by OTUB1-UBE2D2 (OTUB1*).

Table B-1: Primers used in the reprogramming of Yuh1

Primer	Sequencing (5'-3')
pCT-yuh1_f	GGTGGTTCTGCTAGCATGTCGGGTGAAAAC
pCT-yuh1_r	CTTTTGTTCGGATCCTTCCCAGTTCGGGCC
pCT-yuh1_f2	GACAATAGCTCGACGATTGAAGGTAGATACCCATACGA CGTTCCAGACTAC
pCT-yuh1_r2	ATCAGATCTCGAGCTATTACAAGTCCTCTTCAGAAATAA GCT

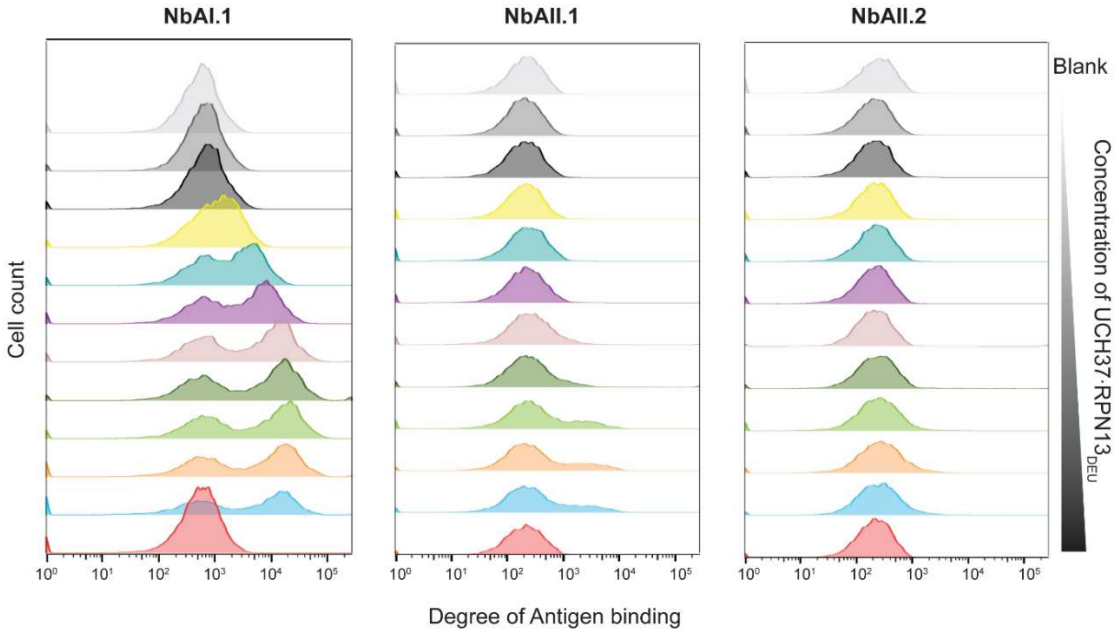

Figure B-4: Titration curve of yeast displayed nanobody with UCH37·RPN13$_{DEU}$

Flow cytometry histograms of yeast displayed NbAI.1 (Left), NbAII.1 (Middle) and NbAII.2 (Right) incubated with biotinylated UCH37·RPN13$_{DEU}$. Cells were labeled with streptavidin conjugated Alexa Fluor® 488.

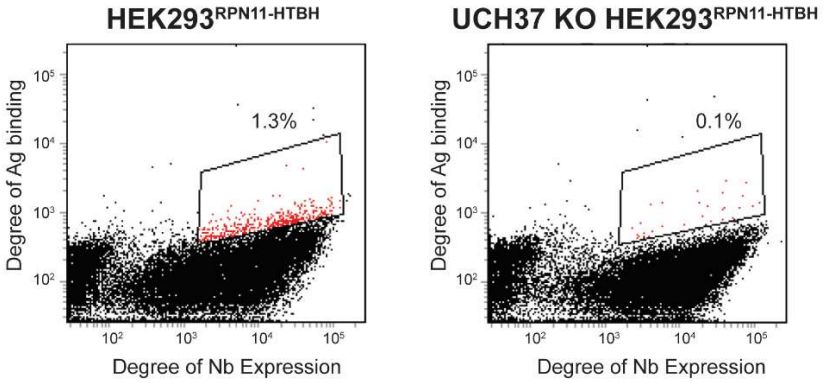

Figure B-5: Validation of the scheme for the affinity maturation of NbAI.1

Flow cytometry dot plot of yeast cells displaying the ensembles of random mutagenized NbAI.1 incubated with WT (Left) and UCH37 KO HEK293FT$^{RPN11\text{-}HTBH}$ (Right). The frame and percentages designate the fraction of yeast cells showing binding towards the antigen.

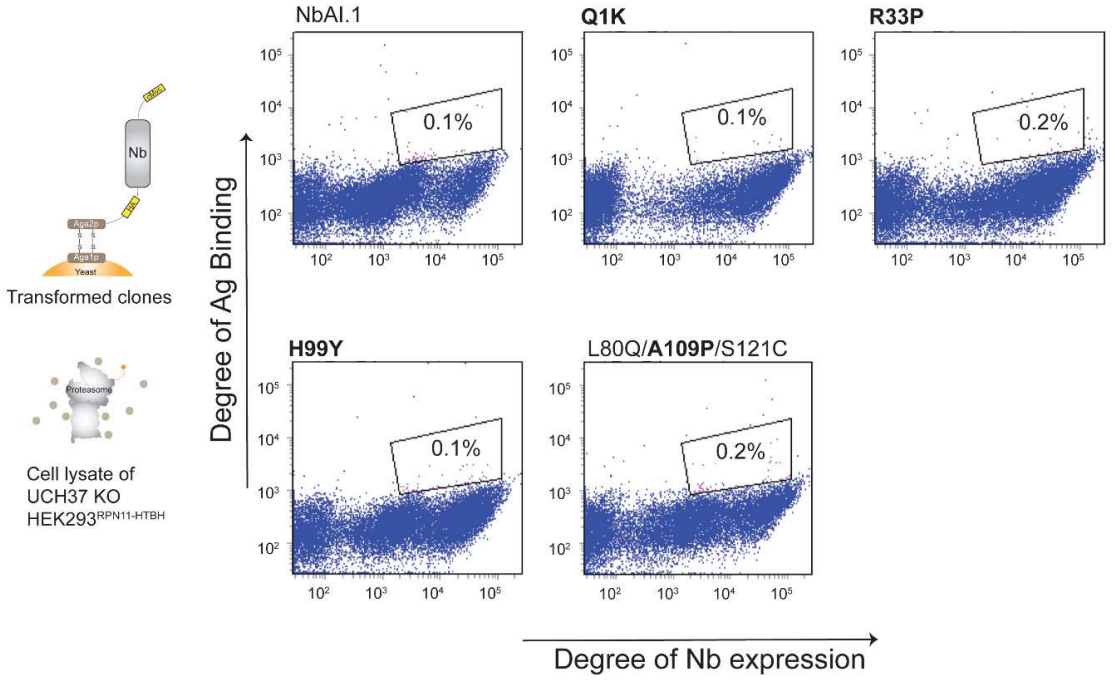

Figure B-6: Investigation of the enriched mutations

Flow cytometry plots showing the lack of binding of the yeast surface displayed-nanobodies containing the enriched mutations towards cell lysate generated from UCH37 KO HEK293$^{RPN11\text{-}HTBH}$. The enriched mutations are highlighted in bold.

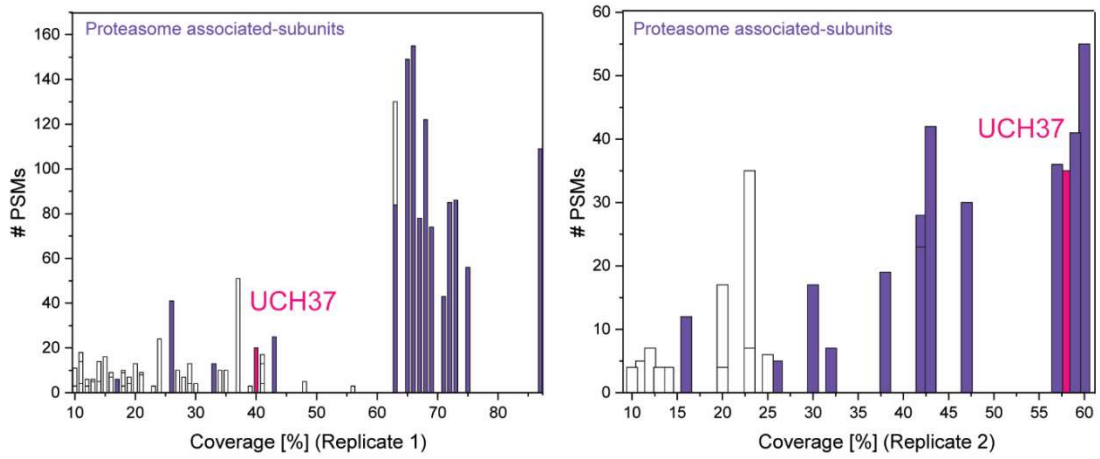

Figure B-7: Specificity assessment of NbAIII.15 through immunoprecipitation-mass spectrometry (IP-MS)

Plots of peptide spectral matches (PSMs) vs coverage for the proteins identified with $p<0.05$ through immunoprecipitation-mass spectrometry. The significance was compared by Student's t test between the protein abundance ratio immunoprecipitated by NbAIII.15 and NbCtrl. Data for UCH37 are highlighted in pink and proteasome-associated subunits are labeled in blue.

Table B-2: NbAIII.15 fragments identified in the limited proteolysis of NbAIII.15 in complex with UCH37·RPN13$_{DEU}$

Measured Mass [M+H]	Calculated Mass [M+H]	Observed amino acid residues
2227	2228	64-84
5108	5108	96-141
6828	6828	1-63
7684	7686	1-141

Only proteolysis fragments generated from NbAIII.15 were listed here. These fragments were detected in two biological replicates. The identities of the proteolysis fragments were determined using ProteinProspector.

Table B-3: Primers for the generation of random mutagenized NbAl.1 library

Primers	Sequence (5'-3')
pCT_Nb_amp_f	GACAATAGCTCGACGATTGAAGGTAGATACCCATACG ACGTTCCAGACTAC
pCT_Nb_amp_r	ATCAGATCTCGAGCTATTACAAGTCCTCTTCAGAAATA AGCT

Table B-4: Primers used in the preparation of library for next generation sequencing

Primers	Sequence (5'-3')
Nbseq_deg_f	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG**NNN**GAC GTTCCAGACTAC
Nbseq_deg_r	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG**NNN**CA AGTCCTCTTCAGA
Nbseq_N502	AATGATACGGCGACCACCGAGATCTACAC***CTCTCTAT***TCG TCGGCAGCG
Nbseq_N503	AATGATACGGCGACCACCGAGATCTACAC***TATCCTCT***TCG TCGGCAGCG
Nbseq_N701	CAAGCAGAAGACGGCATACGAGAT***TCGCCTTA***GTCTCGT GGGCT

Codons used as the unique identified code are highlighted in bold and codons in italic represent the index adaptors for multiplex next generation sequencing.

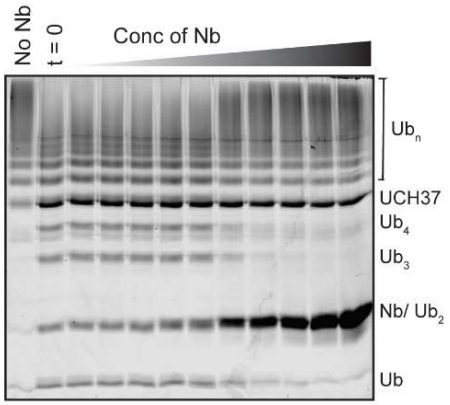

Figure B-8: Investigate the effect of NbAll.1 towards UCH37

Representative gel for the cleavage assay using high molecular weight Lys6-/ Lys48-linked branched ubiquitin chains by UCH37 against a concentration gradient of NbAll.1.

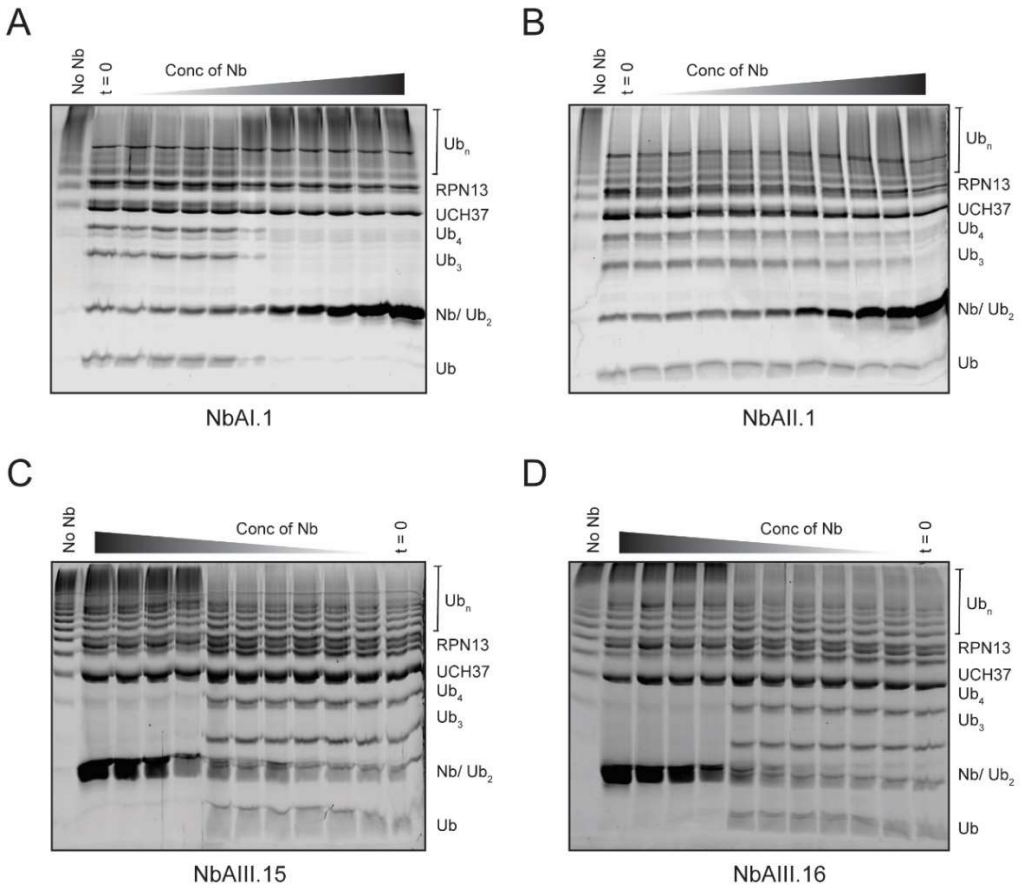

Figure B-9: Investigate the effect of nanobodies towards UCH37·RPN13

Representative gel for the cleavage assay using high molecular weight Lys6-/ Lys48- linked branched ubiquitin chains by UCH37·RPN13 against a concentration gradient of (**A**) NbAI.1, (**B**) NbAII.1, (**C**) NbAIII.15, and (**D**) NbAIII.16.

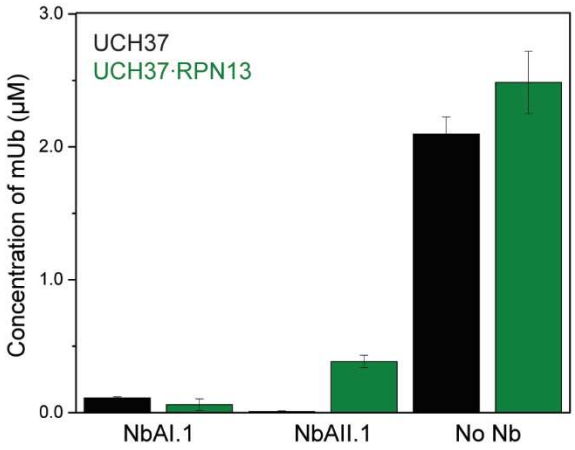

Figure B-10: Comparison of the inhibition effect of NbAII.1 towards UCH37 and UCH37·RPN13

Comparison of the debranching activities of UCH37 and UCH37·RPN13 treated with control, NbAI.1 and NbAII.1. Error bars represent the standard error of two biological replicates.

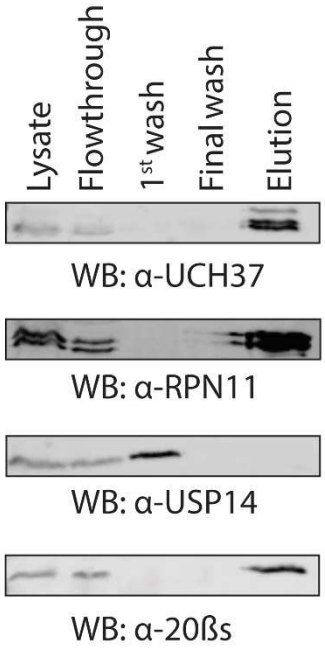

Figure B-11: Purified proteasome-associated UCH37

Western blots of the purified proteasome-associated UCH37 validate the presence of UCH37 and absence of USP14 on the proteasome. Western blot against RPN11 and 20βS, which are the constitutive subunits of the 19S RP and 20S CP confirmed the identity of 26S proteasome complex.

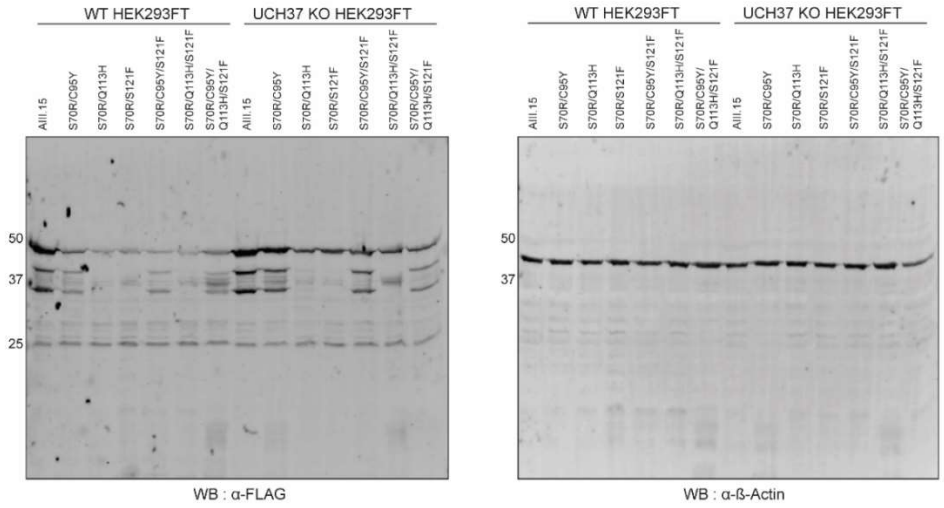

Figure B-12: Conditionally stable NbAIII.15 in WT and UCH37 KO HEK293FT cells

Representative full Western blots showing the protein level (Left) of the nanobodies in the WT and UCH37 KO HEK293FT cells. β-Actin (Right) was used as the loading control.

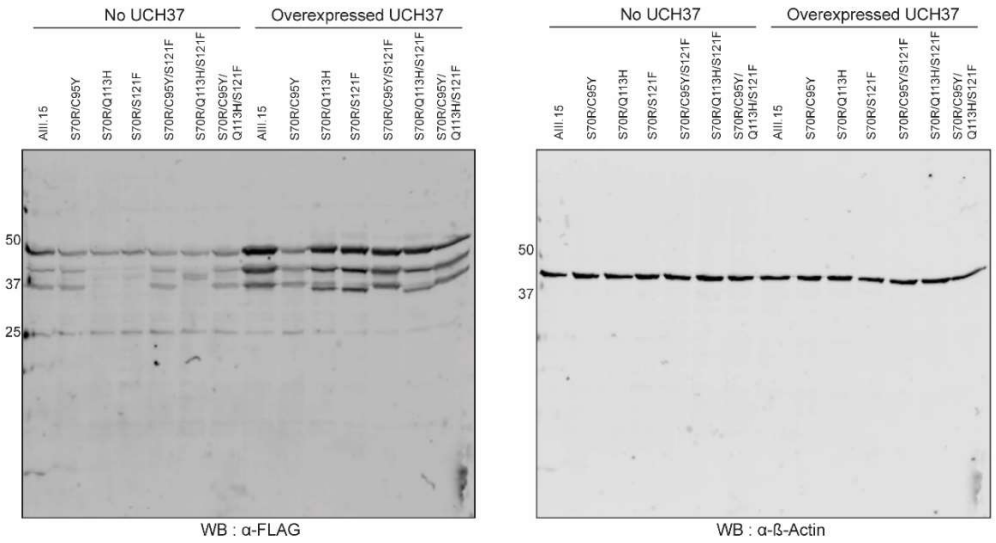

Figure B-13: Conditionally stable NbAIII.15 in UCH37 KO HEK293FT cells without and with ectopically expressed UCH37

Representative full Western blots showing the protein level (Left) of the nanobodies in the UCH37 KO HEK293FT cells without and with overexpressed UCH37. β-Actin (Right) was used as the loading control.

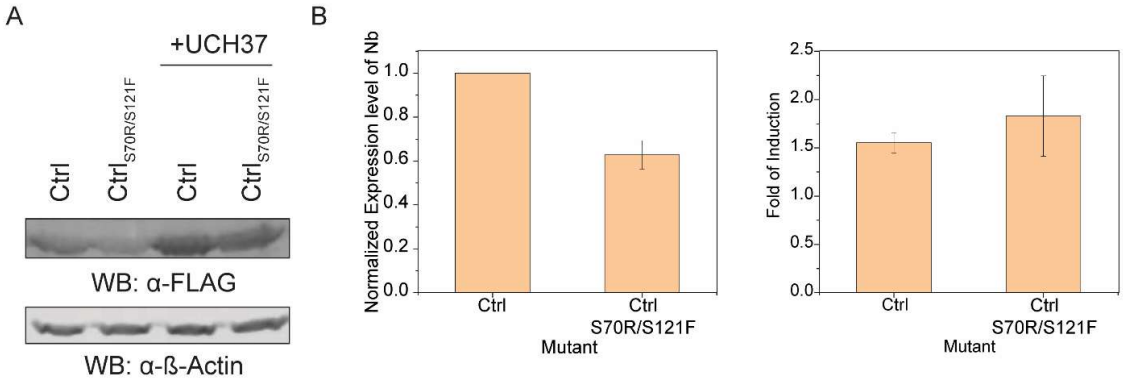

Figure B-14: Conditionally stable NbCtrl in UCH37 KO HEK293FT cells without and with ectopically expressed UCH37

(**A**) Representative Western blot and (**B**) the expression level (Left) and fold of induction (Right) of NbCtrl variants in UCH37 KO HEK293FT cells without and with overexpressed UCH37. Error bars represent the standard error of two biological replicates.

BIBLIOGRAPHY

Ciehanover, A., Hod, Y. & Hershko, A. A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem. Biophys. Res. Commun.* **81**, 1100–1105 (1978).

Hershko, A., Ciechanover, A. & Rose, I. A. Resolution of the ATP dependent proteolytic system from reticulocytes: A component that interacts with ATP. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 3107–3110 (1979).

Hershko, A., Ciechanover, A., Heller, H., Haas, A. L. & Rose, I. A. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 1783–1786 (1980).

Wilkinson, K. D., Urban, M. K. & Haas, A. L. Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *J. Biol. Chem.* **255**, 7529–7532 (1980).

Pickart, C. M. & Eddins, M. J. Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta - Mol. Cell Res.* **1695**, 55–72 (2004).

Reyes-Turcu, F. E., Ventii, K. H. & Wilkinson, K. D. Regulation and Cellular Roles of Ubiquitin-Specific Deubiquitinating Enzymes. *Annu. Rev. Biochem.* **78**, 363–397 (2009).

Xu, P. *et al.* Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. *Cell* **137**, 133–145 (2009).

Crowe, S. O., Rana, A. S. J. B., Deol, K. K., Ge, Y. & Strieter, E. R. Ubiquitin Chain Enrichment Middle-Down Mass Spectrometry Enables Characterization of Branched Ubiquitin Chains in Cellulo. *Anal. Chem.* **89**, 4428–4434 (2017).

Clague, M. J., Heride, C. & Urbé, S. The demographics of the ubiquitin system. *Trends Cell Biol.* **25**, 417–426 (2015).

Kaiser, S. E. *et al.* Protein standard absolute quantification (PSAQ) method for

the measurement of cellular ubiquitin pools. *Nat. Methods* **8**, 691–696 (2011).

Cao, J. & Yan, Q. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Front. Oncol.* **2**, 1–9 (2012).

Pickart, C. M. Targeting of substrates to the 26S proteasome. *FASEB J.* **11**, 1055–1066 (1997).

Uckelmann, M. & Sixma, T. K. Histone ubiquitination in the DNA damage response. *DNA Repair (Amst).* **56**, 92–101 (2017).

Xia, Z. P. *et al.* Direct activation of protein kinases by unanchored polyubiquitin chains. *Nature* **461**, 114–119 (2009).

Komander, D. & Rape, M. The Ubiquitin Code. *Annu. Rev. Biochem.* **81**, 203–229 (2012).

Sloper-Mould, K. E., Jemc, J. C., Pickart, C. M. & Hicke, L. Distinct Functional Surface Regions on Ubiquitin. *J. Biol. Chem.* **276**, 30483–30489 (2001).

Reyes-Turcu, F. E. *et al.* The Ubiquitin Binding Domain ZnF UBP Recognizes the C-Terminal Diglycine Motif of Unanchored Ubiquitin. *Cell* **124**, 1197–1208 (2006).

Varadan, R., Assfalg, M., Raasi, S., Pickart, C. & Fushman, D. Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain. *Mol. Cell* **18**, 687–698 (2005).

Sahtoe, D. D. & Sixma, T. K. Layers of DUB regulation. *Trends Biochem. Sci.* **40**, 456–467 (2015).

Masuda, Y., Kanao, R., Kawai, H., Kukimoto, I. & Masutani, C. Preferential digestion of PCNA-ubiquitin and p53-ubiquitin linkages by USP7 to remove polyubiquitin chains from substrates. *J. Biol. Chem.* **294**, 4177–4187 (2019).

Fejzo, M. S. *et al.* Comprehensive Analysis of 20q13 Genes in Ovarian Cancer

Identifies ADRM1 as Amplification Target. *Genes. Chromosomes Cancer* **47**, 873–883 (2008).

Bilguvar, K. *et al.* Recessive loss of function of the neuronal ubiquitin hydrolase UCHL1 leads to early-onset progressive neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 3489–3494 (2013).

Lander, G. C. *et al.* Complete subunit architecture of the proteasome regulatory particle. *Nature* **482**, 186–191 (2012).

Tanaka, K. The proteasome: Overview of structure and functions. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* **85**, 12–36 (2009).

Paubelle, E. *et al.* Complete remission with bortezomib on plasmocytomas in an end-stage patient with refractory multiple myeloma who failed all other therapies including hematopoietic stem cell transplantation: Possible enhancement of graft-vs-tumor effect [15]. *Leukemia* **19**, 1702–1704 (2005).

Motosugi, R. & Murata, S. Dynamic regulation of proteasome expression. *Front. Mol. Biosci.* **6**, 4–11 (2019).

Voutsadakis, I. A. Proteasome expression and activity in cancer and cancer stem cells. *Tumor Biol.* **39**, 1–17 (2017).

Xu, H. *et al.* The CCAAT box-binding transcription factor NF-Y regulates basal expression of human proteasome genes. *Biochim. Biophys. Acta - Mol. Cell Res.* **1823**, 818–825 (2012).

Pilarsky, C., Wenzig, M., Specht, T., Saeger, H. D. & Grützmann, R. Identification and validation of commonly overexpressed genes in solid tumors by comparison of microarray data. *Neoplasia* **6**, 744–750 (2004).

Boughton, A. J., Krueger, S. & Fushman, D. Branching via K11 and K48 Bestows Ubiquitin Chains with a Unique Interdomain Interface and Enhanced Affinity for Proteasomal Subunit Rpn1. *Structure* **28**, 29-43.e6 (2020).

Radhakrishnan, S. K. *et al.* Transcription Factor Nrf1 Mediates the Proteasome Recovery Pathway after Proteasome Inhibition in Mammalian Cells. *Mol. Cell* **38**, 17–28 (2010).

Tsvetkov, P. *et al.* Compromising the 19S proteasome complex protects cells from reduced flux through the proteasome. *Elife* **4**, 1–22 (2015).

Acosta-Alvear, D. *et al.* Paradoxical resistance of multiple myeloma to proteasome inhibitors by decreased levels of 19S proteasomal subunits. *Elife* **4**, 1–19 (2015).

Xu, D. *et al.* Phosphorylation and activation of ubiquitin-specific protease-14 by Akt regulates the ubiquitin-proteasome system. *Elife* **4**, 1–16 (2015).

Liang, R. Y. *et al.* Rad23 interaction with the proteasome is regulated by phosphorylation of its ubiquitin-like (UbL) domain. *J. Mol. Biol.* **426**, 4049–4060 (2014).

Hemmis, C. W., Heard, S. C. & Hill, C. P. Phosphorylation of Tyr-950 in the proteasome scaffolding protein RPN2 modulates its interaction with the ubiquitin receptor RPN13. *J. Biol. Chem.* **294**, 9659–9665 (2019).

Isasa, M. *et al.* Monoubiquitination of RPN10 Regulates Substrate Recruitment to the Proteasome. *Mol. Cell* **38**, 733–745 (2010).

Besche, H. C. *et al.* Autoubiquitination of the 26S Proteasome on Rpn13 Regulates Breakdown of Ubiquitin Conjugates. *EMBO J.* **33**, 1159–1176 (2014).

Wang, X., Yen, J., Kaiser, P. & Huang, L. Regulation of the 26S proteasome complex during oxidative stress. *Sci. Signal.* **3**, ra88–ra88 (2010).

Aiken, C. T., Kaake, R. M., Wang, X. & Huang, L. Oxidative Stress-Mediated Regulation of Proteasome Complexes. *Mol. Cell. Proteomics* **10**, R110.006924 (2011).

Fabre, B. *et al.* Label-free quantitative proteomics reveals the dynamics of proteasome complexes composition and stoichiometry in a wide range of human cell lines. *J. Proteome Res.* **13**, 3027–3037 (2014).

Fabre, B. *et al.* Subcellular distribution and dynamics of active proteasome complexes unraveled by a workflow combining in vivo complex cross-linking and quantitative proteomics. *Mol. Cell. Proteomics* **12**, 687–699 (2013).

Wang, X. & Huang, L. Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. *Mol. Cell. Proteomics* **7**, 46–57 (2008).

Deol, K. K., Crowe, S. O., Du, J., Bisbee, H. & Strieter, E. R. Proteasome-Bound UCH37 Debranches Ubiquitin Chains to Promote Degradation 1 2. doi:10.1101/2020.02.21.960088.

Lee, B.-H. *et al.* USP14 deubiquitinates proteasome-bound substrates that are ubiquitinated at multiple sites. *Nature* **532**, 398–401 (2016).

de Poot, S. A. H., Tian, G. & Finley, D. Meddling with Fate: The Proteasomal Deubiquitinating Enzymes. *J. Mol. Biol.* **429**, 3525–3545 (2017).

Chadchankar, J. *et al.* Inactive USP14 and inactive UCHL5 cause accumulation of distinct ubiquitinated proteins in mammalian cells. *PLoS One* **14**, 1–23 (2019).

Liu, B. *et al.* Proteome-wide analysis of USP14 substrates revealed its role in hepatosteatosis via stabilization of FASN. *Nat. Commun.* **9**, 1–12 (2018).

Lee, B. H. *et al.* Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* **467**, 179–184 (2010).

Zhao, B. *et al.* Protein engineering in the ubiquitin system: Tools for discovery and beyond. *Pharmacol. Rev.* **72**, 380–413 (2020).

Chen, R. Enzyme engineering: Rational redesign versus directed evolution. *Trends Biotechnol.* **19**, 13–14 (2001).

Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **10**, 866–876 (2009).

Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nat. Publ. Gr.* **16**, 379–394 (2015).

Smith, G. P. & Petrenko, V. A. Phage display. *Chem. Rev.* **97**, 391–410 (1997).

Daugherty, P. S. Protein engineering with bacterial display. *Curr. Opin. Struct. Biol.* **17**, 474–480 (2007).

Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **15**, 553–557 (1997).

Akamatsu, Y., Pakabunto, K., Xu, Z., Zhang, Y. & Tsurushita, N. Whole IgG surface display on mammalian cells: Application to isolation of neutralizing chicken monoclonal anti-IL-12 antibodies. *J. Immunol. Methods* **327**, 40–52 (2007).

Zahnd, C., Amstutz, P. & Plückthun, A. Ribosome display: Selecting and evolving proteins in vitro that specifically bind to a target. *Nat. Methods* **4**, 269–279 (2007).

Wilson, D. S., Keefe, A. D. & Szostak, J. W. The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3750–3755 (2001).

Ghadessy, F. J., Ong, J. L. & Holliger, P. *Directed evolution of polymerase function by compartmentalized self-replication*. www.pnas.orgcgidoi10.1073pnas.071052198.

Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci.* **108**, 11399–11404 (2011).

Glanville, J. *et al.* Deep sequencing in library selection projects: What insight does it bring? *Curr. Opin. Struct. Biol.* **33**, 146–160 (2015).

Wrenbeck, E. E., Faber, M. S. & Whitehead, T. A. Deep sequencing methods for protein engineering and design. *Curr. Opin. Struct. Biol.* **45**, 36–44 (2017).

Portnoff, A. D., Stephens, E. A., Varner, J. D. & DeLisa, M. P. Ubiquibodies, synthetic E3 ubiquitin ligases endowed with unnatural substrate specificity for targeted protein silencing. *J. Biol. Chem.* **289**, 7844–7855 (2014).

Yau, R. G. *et al.* Assembly and Function of Heterotypic Ubiquitin Chains in Cell-Cycle and Protein Quality Control. *Cell* **171**, 918-933.e20 (2017).

Michel, M. A., Swatek, K. N., Hospenthal, M. K. & Komander, D. Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. *Mol. Cell* **68**, 233-246.e5 (2017).

Ernst, A. *et al.* A Strategy for Modulation of Enzymes in the Ubiquitin System. *Science (80-.).* **339**, 590–595 (2013).

Manczyk, N. *et al.* Structural and functional characterization of a ubiquitin variant engineered for tight and specific binding to an alpha-helical ubiquitin interacting motif. *Protein Sci.* **26**, 1060–1069 (2017).

Wyler, E. *et al.* Inhibition of NF-κB activation with designed ankyrin-repeat proteins targeting the ubiquitin-binding/oligomerization domain of NEMO. *Protein Sci.* **16**, 2013–2022 (2007).

Grubisha, O. *et al.* DARPin-Assisted Crystallography of the CC2-LZ Domain of NEMO Reveals a Coupling between Dimerization and Ubiquitin Binding. *J. Mol. Biol.* **395**, 89–104 (2010).

Chang, L. H. & Strieter, E. R. Reprogramming a Deubiquitinase into a Transamidase. *ACS Chem. Biol.* **13**, 2808–2818 (2018).

Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).

Avram Hershko, A. C. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425 (1998).

Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).

Komander, D., Clague, M. J. & Urbé, S. Breaking the chains: Structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**, 550–563 (2009).

Mevissen, T. E. T. & Komander, D. Mechanisms of Deubiquitinase Specificity and Regulation. *Annu. Rev. Biochem.* **86**, 159–192 (2017).

Dong, K. C. *et al.* Preparation of distinct ubiquitin chain reagents of high purity and yield. *Structure* **19**, 1053–1063 (2011).

Martinez-Fonts, K. & Matouschek, A. A Rapid and Versatile Method for Generating Proteins with Defined Ubiquitin Chains. *Biochemistry* **55**, 1898–1908 (2016).

Martinez-Fonts, K. *et al.* The proteasome 19S cap and its ubiquitin receptors provide a versatile recognition platform for substrates. *Nat. Commun.* **11**, (2020).

Saeki, Y., Isono, E. & Toh-E, A. Preparation of ubiquitinated substrates by the PY motif-insertion method for monitoring 26S proteasome activity. *Methods Enzymol.* **399**, 215–227 (2005).

Chatterjee, C., McGinty, R. K., Pellois, J. P. & Muir, T. W. Auxiliary-mediated site-specific peptide ubiquitylation. *Angew. Chemie - Int. Ed.* **46**, 2814–2818 (2007).

McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G. & Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453**, 812–816 (2008).

Johnston, S. C., Riddle, S. M., Cohen, R. E. & Hill, C. P. Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J.* **18**, 3877–3887 (1999).

Valkevich, E. M. *et al.* Forging isopeptide bonds using thiol-ene chemistry: Site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases. *J. Am. Chem. Soc.* **134**, 6916–6919 (2012).

Trang, V. H., Rodgers, M. L., Boyle, K. J., Hoskins, A. A. & Strieter, E. R. Chemoenzymatic synthesis of bifunctional polyubiquitin substrates for monitoring ubiquitin chain remodeling. *ChemBioChem* **15**, 1563–1568 (2014).

Yu, H. A. *et al.* Characterization of ubiquitin C-terminal hydrolase 1 (YUH1) from Saccharomyces cerevisiae expressed in recombinant Escherichia coli. *Protein Expr. Purif.* **56**, 20–26 (2007).

Ritorto, M. S. *et al.* Screening of DUB activity and specificity by MALDI-TOF mass spectrometry. *Nat. Commun.* **5**, (2014).

Khersonsky, O. & Tawfik, D. S. Enzyme Promiscuity : A Mechanistic and Evolutionary Perspective. (2010) doi:10.1146/annurev-biochem-030409-143718.

Copley, S. D. Shining a light on enzyme promiscuity. *Curr. Opin. Struct. Biol.* **47**, 167–175 (2017).

De Bie, P. & Ciechanover, A. Ubiquitination of E3 ligases: Self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms. *Cell Death Differ.* **18**, 1393–1402 (2011).

Boder, E. T. & Wittrup, K. D. Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol.* **328**, 430–444 (2000).

Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. **1**, 755–769 (2007).

Chen, T. F., De Picciotto, S., Hackel, B. J. & Wittrup, K. D. *Engineering fibronectin-based binding proteins by yeast surface display*. *Methods in Enzymology* vol. 523 (Elsevier Inc., 2013).

Jan Deniau, A. H., Subileau, M. & Dubreucq, E. Characterization and Reshaping of a Large and Hydrophobic Nucleophile Pocket in Lipases/Acyltransferases. *ChemBioChem* **19**, 1839–1844 (2018).

Harper, S. *et al.* Structure and catalytic regulatory function of ubiquitin specific protease 11 N-terminal and ubiquitin-like domains. *Biochemistry* **53**, 2966–2978 (2014).

Crowe, S. O. *et al.* Subunit-Specific Labeling of Ubiquitin Chains by Using Sortase: Insights into the Selectivity of Deubiquitinases. *ChemBioChem* 1525–1531 (2016) doi:10.1002/cbic.201600276.

Ye, Y. *et al.* Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21. *EMBO Rep.* **12**, 350–357 (2011).

Edelmann, M. J. *et al.* Structural basis and specificity of human otubain 1-mediated deubiquitination. *Biochem. J.* **418**, 379–390 (2009).

Altun, M. *et al.* The human otubain2-ubiquitin structure provides insights into the cleavage specificity of poly-ubiquitin-linkages. *PLoS One* **10**, 1–15 (2015).

Mevissen, T. E. T. *et al.* OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell* **154**, 169–184 (2013).

Wiener, R. *et al.* E2 ubiquitin-conjugating enzymes regulate the deubiquitinating activity of OTUB1. *Nat. Struct. Mol. Biol.* **20**, 1033–1039 (2013).

Hospenthal, M. K., Freund, S. M. V. & Komander, D. Assembly, analysis and architecture of atypical ubiquitin chains. *Nat. Struct. Mol. Biol.* **20**, 555–565 (2013).

Wiener, R., Zhang, X., Wang, T. & Wolberger, C. The mechanism of OTUB1-mediated inhibition of ubiquitination. *Nature* **483**, 618–622 (2012).

Schulman, B. A. & Wade Harper, J. Ubiquitin-like protein activation by E1 enzymes: The apex for downstream signalling pathways. *Nat. Rev. Mol. Cell Biol.* **10**, 319–331 (2009).

Schaefer, J. B. & Morgan, D. O. Protein-linked ubiquitin chain structure restricts activity of deubiquitinating enzymes. *J. Biol. Chem.* **286**, 45186–45196 (2011).

Voges, D., Zwickl, P. & Baumeister, W. The 26S Proteasome: A molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* **68**, 1015–1068 (1999).

Grice, G. L. & Nathan, J. A. The recognition of ubiquitinated proteins by the proteasome. *Cell. Mol. Life Sci.* **73**, 3497–3506 (2016).

Qiu, X. B. *et al.* hRpn13/ADRM1/GP110 is a novel proteasome subunit that binds the deubiquitinating enzyme, UCH37. *EMBO J.* **25**, 5742–5753 (2006).

Hamazaki, J. *et al.* A novel proteasome interacting protein recruits the deubiquitinating enzyme UCH37 to 26S proteasomes. *EMBO J.* **25**, 4524–4536 (2006).

Yao, T. *et al.* Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1. *Nat. Cell Biol.* **8**, 994–1002 (2006).

Nishio, K. *et al.* Crystal structure of the de-ubiquitinating enzyme UCH37 (human UCH-L5) catalytic domain. *Biochem. Biophys. Res. Commun.* **390**, 855–860 (2009).

Burgie, S. E., Bingman, C. A., Soni, A. B. & Phillips, G. N. Structural characterization of human Uch37. *Proteins Struct. Funct. Bioinforma.* **80**, 649–654 (2012).

Sahtoe, D. D. *et al.* Mechanism of UCH-L5 Activation and Inhibition by DEUBAD Domains in RPN13 and INO80G. *Mol. Cell* **57**, 887–900 (2015).

VanderLinden, R. T. *et al.* Structural Basis for the Activation and Inhibition of the UCH37 Deubiquitylase. *Mol. Cell* **57**, 901–911 (2015).

D'Arcy, P. *et al.* Inhibition of proteasome deubiquitinating activity as a new cancer therapy. *Nat. Med.* **17**, 1636–1640 (2011).

Lu, X. *et al.* Structure of the Rpn13-Rpn2 complex provides insights for Rpn13 and Uch37 as anticancer targets. *Nat. Commun.* **8**, 15540 (2017).

Aufderheide, A. *et al.* Structural characterization of the interaction of Ubp6 with the 26S proteasome. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8626–8631 (2015).

Stocks, M. Intrabodies as drug discovery tools and therapeutics. *Curr. Opin. Chem. Biol.* **9**, 359–365 (2005).

Gupta, A. *et al.* Facile target validation in an animal model with intracellularly expressed monobodies. *Nat. Chem. Biol.* 1–6 (2018) doi:10.1038/s41589-018-0099-z.

Schumacher, D., Helma, J., Schneider, A. F. L., Leonhardt, H. & Hackenberger, C. P. R. Nanobodies : Chemical Functionalization Strategies and Intracellular Applications Angewandte. 2314–2333 doi:10.1002/anie.201708459.

Dmitriev, O. Y., Lutsenko, S. & Muyldermans, S. Nanobodies as probes for protein dynamics in vitro and in cells. *J. Biol. Chem.* **291**, 3767–3775 (2016).

Mcmahon, C. *et al.* Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, (2018).

Zimmermann, I. *et al.* Synthetic single domain antibodies for the conformational trapping of membrane proteins. *Elife* **3**, 1–32 (2018).

Jiao, L. *et al.* Mechanism of the Rpn13-induced activation of Uch37. *Protein Cell* **5**, 616–630 (2014).

Gautier, A. & Hinner, M. J. Site-specific biotinylation of purified proteins using BirA. *Site-Specific Protein Labeling Methods Protoc.* 171–184 (2015) doi:10.1007/978-1-4939-2272-7.

Ackerman, M., Levary, D., Tobon, G., Hackel, B. & Orcutt, K. D. Highly Avid Magnetic Bead Capture: An Efficient Selection Method. *Artif. Cells Blood Substitutes Immobil. Biotechnol. Cells Blood Substit Immobi* **25**, 774–783 (2010).

Hunter, S. A. & Cochran, J. R. *Cell-Binding Assays for Determining the Affinity of Protein–Protein Interactions: Technologies and Considerations*. *Methods in Enzymology* vol. 580 (Elsevier Inc., 2016).

Wang, X. *et al.* Mass Spectrometric Characterization of the Affinity-Purified Human 26S Proteasome Complex †. (2007) doi:10.1021/bi061994u.

Kuo, C.-L., Collins, G. A. & Goldberg, A. L. Methods to Rapidly Prepare Mammalian 26S Proteasomes for Biochemical Analysis. *Methods Mol. Biol.* 277–288 (2018).

Tillotson, B. J., De Larrinoa, I. F., Skinner, C. A., Klavas, D. M. & Shusta, E. V. Antibody affinity maturation using yeast display with detergent-solubilized membrane proteins as antigen sources. *Protein Eng. Des. Sel.* **26**, 101–112 (2013).

Mason, D. M. *et al.* High-throughput antibody engineering in mammalian cells by CRISPR/Cas9-mediated homology-directed mutagenesis. *Nucleic Acids Res.* **46**, 7436–7449 (2018).

Deshpande, I. *et al.* Smoothened stimulation by membrane sterols drives Hedgehog pathway activity. *Nature* **571**, 284–288 (2019).

Karlsson, R., Katsamba, P. S., Nordin, H., Pol, E. & Myszka, D. G. Analyzing a kinetic titration series using affinity biosensors. *Anal. Biochem.* **349**, 136–147 (2006).

Suh, M. J., Pourshahian, S. & Limbach, P. A. Developing Limited Proteolysis and Mass Spectrometry for the Characterization of Ribosome Topography. *J. Am. Soc. Mass Spectrom.* **18**, 1304–1317 (2007).

ProteinProspector. http://prospector.ucsf.edu/prospector/mshome.htm.

Los, G. V. *et al.* HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008).

Yao, T. *et al.* Distinct Modes of Regulation of the Uch37 Deubiquitinating Enzyme in the Proteasome and in the Ino80 Chromatin-Remodeling Complex. *Mol. Cell* **31**, 909–917 (2008).

Uchański, T. *et al.* An improved yeast surface display platform for the screening of nanobody immune libraries. *Sci. Rep.* **9**, 1–12 (2019).

Wang, Z., Mathias, A., Stavrou, S. & Neville, D. M. A new yeast display vector permitting free scFv amino termini can augment ligand binding affinities. *Protein Eng. Des. Sel.* **18**, 337–343 (2005).

Andrews, S. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ doi:10.1016/0038-0717(73)90093-X.

Magoč, T. & Salzberg, S. L. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).

BBMap download | SourceForge.net. https://sourceforge.net/projects/bbmap/.

Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**, (2016).

Livneh, I., Cohen-Kaplan, V., Cohen-Rosenzweig, C., Avni, N. & Ciechanover, A. The life cycle of the 26S proteasome: From birth, through regulation and function, and onto its death. *Cell Res.* **26**, 869–885 (2016).

Collins, G. A. & Goldberg, A. L. The Logic of the 26S Proteasome. *Cell* **169**, 792–806 (2017).

Hamazaki, J., Hirayama, S. & Murata, S. Redundant Roles of Rpn10 and Rpn13 in Recognition of Ubiquitinated Proteins and Cellular Homeostasis. *PLoS Genet.* **11**, 1–20 (2015).

Koulich, E., Li, X. & N. DeMartino, G. Relative Structural and Functional Roles of Multiple Deubiquitylating Proteins Associated with Mammalian 26S Proteasome. *Mol. Biol. Cell* **19**, 1072–1082 (2008).

Berko, D. *et al.* Inherent asymmetry in the 26S proteasome is defined by the ubiquitin receptor RPN13. *J. Biol. Chem.* **289**, 5609–5618 (2014).

Lundgren, J., Masson, P., Realini, C. A. & Young, P. Use of RNA Interference and Complementation To Study the Function of the Drosophila and Human 26S Proteasome Subunit S13. *Mol. Cell. Biol.* **23**, 5320–5330 (2003).

Traenkle, B. & Rothbauer, U. Under the microscope: Single-domain antibodies for live-cell imaging and super-resolution microscopy. *Front. Immunol.* **8**, 1–8 (2017).

Keller, B. M. *et al.* Chromobodies to quantify changes of endogenous protein concentration in living cells. *Mol. Cell. Proteomics* **17**, 2518–2533 (2018).

Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–898 (2018).

Tachie-Menson, T. *et al.* Characterisation of the biochemical and cellular roles of native and pathogenic amelogenesis imperfecta mutants of FAM83H. *Cell. Signal.* **72**, 109632 (2020).

Tang, J. C. Y. *et al.* Detection and manipulation of live antigen-expressing cells using conditionally stable nanobodies. *Elife* **5**, 1–27 (2016).

Al-Lazikani, B., Lesk, A. M. & Chothia, C. Standard conformations for the canonical structures of immunoglobulins. *J. Mol. Biol.* **273**, 927–948 (1997).

Ariotti, N. *et al.* Ultrastructural localisation of protein interactions using conditionally stable nanobodies. *PLoS Biol.* **16**, 1–11 (2018).

Meerbrey, K. L. *et al.* The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3665–3670 (2011).

Leestemaker, Y. *et al.* Proteasome Activation by Small Molecules. *Cell Chem. Biol.* **24**, 725-736.e7 (2017).

Schipper-Krom, S. *et al.* Visualizing Proteasome Activity and Intracellular Localization Using Fluorescent Proteins and Activity-Based Probes. *Front. Mol. Biosci.* **6**, 1–18 (2019).

Buchfellner, A. *et al.* A new nanobody-based biosensor to study endogenous PARP1 in vitro and in live human cells. *PLoS One* **11**, 1–23 (2016).

Stadler, C. *et al.* Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nat. Methods* **10**, 315–323 (2013).

Albert, S. *et al.* Direct visualization of degradation microcompartments at the ER membrane. *Proc. Natl. Acad. Sci.* 1–12 (2019) doi:10.1073/pnas.1905641117.

Nakamura, N. & Hirose, S. Regulation of Mitochondrial Morphology by USP30, a Deubiquitinating Enzyme Present in the Mitochondrial Outer Membrane. *Mol. Biol. Cell* **19**, 1903–1911 (2008).

Bingol, B. *et al.* The mitochondrial deubiquitinase USP30 opposes parkin-mediated mitophagy. *Nature* **510**, 370–375 (2014).

Cunningham, C. N. *et al.* USP30 and parkin homeostatically regulate atypical ubiquitin chains on mitochondria. *Nat. Cell Biol.* **17**, 160–169 (2015).

Gersch, M. *et al.* Mechanism and regulation of the Lys6-selective deubiquitinase USP30. *Nat. Struct. Mol. Biol.* **24**, 920–930 (2017).

Sato, Y. *et al.* Structural basis for specific cleavage of Lys6-linked polyubiquitin chains by USP30. *Nat. Struct. Mol. Biol.* **24**, 911–919 (2017).

Dunbar, J. *et al.* SAbPred: a structure-based antibody prediction server. *Nucleic Acids Res.* **44**, W474–W478 (2016).

Sircar, A., Sanni, K. A., Shi, J. & Gray, J. J. Analysis and Modeling of the Variable Region of Camelid Single-Domain Antibodies. *J. Immunol.* **186**, 6357–6367 (2011).

Mitchell, L. S. & Colwell, L. J. Comparative analysis of nanobody sequence and structure data. *Proteins Struct. Funct. Bioinforma.* **86**, 697–706 (2018).

Li, X. *et al.* Comparative analysis of immune repertoires between bactrian Camel's conventional and heavy-chain antibodies. *PLoS One* **11**, 1–15 (2016).

Xu, J. L. & Davis, M. M. Diversity in the CDR3 Region of V. *Immunity* **13**, 37–45 (2000).

Hackel, B. J., Kapila, A. & Wittrup, K. D. Picomolar Affinity Fibronectin Domains Engineered Utilizing Loop Length Diversity, Recursive Mutagenesis, and Loop Shuffling. (2008) doi:10.1016/j.jmb.2008.06.051.

Ono, A., Matsuda, A., Zhao, J. & Santi, D. V. The synthesis of blocked triplet-phosphoramidites and their use in mutagenesis. *Nucleic Acids Res.* **23**, 4677–4682 (1995).